

STIC-Biotech/ChemLib

From: Mertz, Prema  
Sent: Tuesday, June 04, 2002 12:09 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/828,217

Please search SEQ ID NO:1 with protein databases.

Thanks  
Prema Mertz, Ph.D.  
Art Unit 1646  
Mailbox 10C-01  
Crystal Mall 1, Room 10E-01  
United States Patent & Trademark Office  
# (703) 308-4229

CRFE

Searcher: AB Wesen - July  
Phone: 308-4501/4508  
Location: Biotech Lab.  
Date Picked Up: 6/4/02  
Date Completed: 6/5/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: 3 min  
Online time: 2 min

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST(where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: AB5502  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THIS PAGE BLANK (USPTO)



Db 575 GAADLQNRRIARLEAESSGLTRRIQESAAEVALRRRDELRRTAAS----AAQDVS 627

## RESULT 2

ID 09K6X4 PRELIMINARY; PRT: 461 AA.

AC 09K6X4; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 GN CEL1. WALL-BINDING PROTEIN.

OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]

SEQUENCE FROM N.A.

STRATIN-C-125 / JCM 9153;  
 MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL; AP001519; BAB07319.1; -

DR MEROPS; M37.0PM; -

DR InterPro: IPR002886; Peptidase\_M37.

DR Pfam: PF01551; Peptidase\_M37; 1.

KW Complete proteome.  
 SQ SEQUENCE 461 AA; 50372 MW; 2918480CD67AF3F CRC64;

Query Match 12.1%; Score 107.5; DB 16; Length 461;  
 Best Local Similarity 26.1%; Pred. No. 0.44; Indels 5; Gaps 2;  
 Matches 31; Conservative 28; Mismatches 55;

QY 33 IIVLGVLIIITIKANSEACRDGLRAVMECRVTHLLOELTEAKGFQDVEAQAATCN 92  
 Db 5 ISVVAAGLITFSLFSSQSIDA-KANSSLQNGISDVQKREKQEKTELE---- 59

QY 93 HTYMALMASIDAEKAGQKVEELEGITTLNKLQDASAEVERLRRENOVLSYRIADK 151  
 Db 60 KEVEKELGDTAEIERLDEKVEETSGRIQEKREIEVEQAIEELKQIEILERRIAR 118

## RESULT 3

ID 09KA43 PRELIMINARY; PRT: 143 AA.

AC 09KA43; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 GN FLAGELLAR PROTEIN REQUIRED FOR FLAGELLAR FORMATION.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=86665;  
 RN [1]

SEQUENCE FROM N.A.

STRATIN-C-125 / JCM 9153;  
 MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL; AP001515; BAB06166.1; -  
 KW Flagella; Complete proteome.

SQ SEQUENCE 143 AA; 16386 MW; 613446064B1C5402 CRC64;

Query Match 12.0%; Score 107; DB 16; Length 143;  
 Best Local Similarity 22.1%; Pred. No. 0.13; Indels 26; Gaps 4;  
 Matches 31; Conservative 34; Mismatches 49;

QY 23 LIGGILVLLIIVLIGVLIITFKANSEACRDGLRAVMEC-----RNVTHLLOEL 74  
 Db 6 LVNMLILVLIIVLVGVAVLIFVNFNNEDEQDEPTIDELINQSYTEETITNLISDF 65

QY 75 TEA-----QKGEQVEAQAATCNHYMALMASIDAEKAGQKVEELEGITTLNKH 126  
 Db 66 VRAEFLIHDVNRNALQEVQRDEQVNNIIRSLAGMDASQSGADGIEKLEAQ----- 118

QY 127 LODASAEVERLRRENOVLSY 146  
 Db 119 LOD--DINALMQEGSVYKI 135

SEQUENCE FROM N.A.

MEDLINE=20014706; PubMed=10545595;

RA Oka A., Tamiya G., Tomizawa M., Ota M., Katsuyama Y., Makino S.,  
 Shima T., Yoshitome M., Lizuka M., Sasaki Y., Iwashita K.,  
 Kawakubo Y., Sugai J., Ozawa A., Ohkido M., Kimura M., Bahram S.,  
 Inoko H.;

RT "Association analysis using refined microsatellite markers localizes a  
 susceptibility locus for psoriasis vulgaris within a 11kb segment  
 telomeric to the HLA-C gene."

RL Hum. Mol. Genet. 8:2165-2170(1999).

DR EMBL; AB029331; BAA81890.1; -

DR SQUENCE 756 AA; 86118 MW; 88FD5F858EF07601 CRC64;

QY 54 RDLGLRAVME-----CRNVTHL---QOELTEAKGFQDVEAQAATC-----NHTYMA 97  
 Db 258 RDLHRTAFELLYRVOSLTHLLOELTRKVPDSLSPEPTRKQSLNWRKRVRA 317

QY 98 LMASIDAEKAGQKVEELEGITTLNKH-----LODASAEVERLRRENOV 143  
 Db 318 LMVQLAKQELHSDSVKQLQGVASLOEKVTSQSQDAILQRLQDKAAEVERRGANG 377

QY 144 LSVRIA-----DKKYRSSQDSSAANPOLIYVLGLSA 177  
 Db 378 LOELSRAGEARRW--DOOTASAEQRLVYNAVVS 412

SEQUENCE FROM N.A.

MEDLINE=20014706; PubMed=10545595;

RA Oka A., Tamiya G., Tomizawa M., Ota M., Katsuyama Y., Makino S.,  
 Shima T., Yoshitome M., Lizuka M., Sasaki Y., Iwashita K.,  
 Kawakubo Y., Sugai J., Ozawa A., Ohkido M., Kimura M., Bahram S.,  
 Inoko H.;

RT "Association analysis using refined microsatellite markers localizes a  
 susceptibility locus for psoriasis vulgaris within a 11kb segment  
 telomeric to the HLA-C gene."

RL Hum. Mol. Genet. 8:2165-2170(1999).

DR EMBL; AB029331; BAA81890.1; -

DR SQUENCE 756 AA; 86118 MW; 88FD5F858EF07601 CRC64;

Query Match 11.8%; Score 104.5; DB 4; Length 756;  
 Best Local Similarity 24.7%; Pred. No. 1.4; Indels 37; Gaps 7;  
 Matches 39; Conservative 34; Mismatches 48;

QY 54 RDLGLRAVME-----CRNVTHL---QOELTEAKGFQDVEAQAATC-----NHTYMA 97  
 Db 258 RDLHRTAFELLYRVOSLTHLLOELTRKVPDSLSPEPTRKQSLNWRKRVRA 317

QY 98 LMASIDAEKAGQKVEELEGITTLNKH-----LODASAEVERLRRENOV 143  
 Db 318 LMVQLAKQELHSDSVKQLQGVASLOEKVTSQSQDAILQRLQDKAAEVERRGANG 377

QY 144 LSVRIA-----DKKYRSSQDSSAANPOLIYVLGLSA 177  
 Db 378 LOELSRAGEARRW--DOOTASAEQRLVYNAVVS 412

SEQUENCE FROM N.A.

MEDLINE=20014706; PubMed=10545595;

RA Oka A., Tamiya G., Tomizawa M., Ota M., Katsuyama Y., Makino S.,  
 Shima T., Yoshitome M., Lizuka M., Sasaki Y., Iwashita K.,  
 Kawakubo Y., Sugai J., Ozawa A., Ohkido M., Kimura M., Bahram S.,  
 Inoko H.;

RT "Association analysis using refined microsatellite markers localizes a  
 susceptibility locus for psoriasis vulgaris within a 11kb segment  
 telomeric to the HLA-C gene."

RL Hum. Mol. Genet. 8:2165-2170(1999).

DR EMBL; AB029331; BAA81890.1; -

DR SQUENCE 756 AA; 86118 MW; 88FD5F858EF07601 CRC64;

GN HCR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20347693; PubMed-10888604;  
 RA Asunalahti K., Laitinen T., Itkonen-Valtjus R., Lokki M.-L.,  
 RA Suomeala S., Stellan E., Saarialho-Kere U., Kere J.;  
 RT "A candidate gene for psoriasis near HLA-C, HCR (P98), is highly  
 RT polymorphic with a disease-associated susceptibility allele."  
 RL Hum. Mol. Genet. 9:1533-1542(2000).  
 DR EMBL: AF216493; AAF74221.1;  
 SO SEQUENCE 756 AA; 86058 MW; 8EAD03358B62DEB4 CRC64;

Query Match 11.8%; Score 104.5; DB 4; Length 756;  
 Best Local Similarity 24.7%; Pred. No. 1.4;  
 Matches 39; Conservative 34; Mismatches 48; Indels 37; Gaps 7;  
 QY 54 RGLRAVME-----CRNTHLL---QQLTEAQGFQDVEAQ-AAATC-----NHTVMA 97  
 DB 258 RDSLHATAELQVROSLTHIALQEEELTRKVPDSLEPFRKQSLNMRREKVFA 317  
 QY 98 LMASLDAEKAQGGKVEELEGETTTLNKH-----LQDASAEVERLRRENOV 143  
 DB 318 LMVQLAKQLEHSDSVKOLKGVASLOEYVTSOSQEQAILQSLQDKAAEVEVERMGAKG 377  
 QY 144 LSVRIA---DKKYPSSODSSSAAPOLLIVLGLSA 177  
 DB 378 LQLELSRAQEARRW---QQQTASAEQRLVYNAVSS 412

RESULT 6  
 Q9Y6M1 PRELIMINARY: PRT; 756 AA.  
 AC Q9Y6M1;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HCR PROTEIN (TRICHOVALIN HOMOLOGUE).  
 GN HCR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BLOOD;  
 RA Oka A., Tamiya G., Makino S., Tomizawa M., Yamagata T., Shilina T.,  
 RA Watanabe K., Yamazaki M., Tashiro H., Kimura M., Inoko H.;  
 RT "HCR-a-helix coiled-coil rod homologue."  
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Shilina S., Tamiya G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."  
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB028343; BAA82158.1;  
 DR EMBL: AF000509; BAB63313.1;  
 SO SEQUENCE 756 AA; 85951 MW; 3D109AF0EEBA9F5 CRC64;

Query Match 11.8%; Score 104.5; DB 4; Length 756;  
 Best Local Similarity 25.2%; Pred. No. 1.4;  
 Matches 40; Conservative 32; Mismatches 48; Indels 39; Gaps 7;  
 QY 54 RGLRAVME-----CRNTHLL---QQLTEAQGFQDVEAQ-AAATC-----NHTVMA 97

DB 258 RDSLHATAELQVROSLTHIALQEEELTRKVPDSLEPFRKQSLNMRREKVFA 317  
 QY 98 LMASLDAEKAQGGKVEELEGETTTLNKH-----LQDASAEVERLRRENOV 143  
 DB 318 LMVQLAKQLEHSDSVKOLKGVASLOEYVTSOSQEQAILQSLQDKAAEVEVERMGAKG 377  
 QY 144 LSVRIA---ADKKYPSSODSSSAAPOLLIVLGLSA 177  
 DB 378 LQLELSRAQEARRW---QQQTASAEQRLVYNAVSS 412

RESULT 7  
 Q9NXK3 PRELIMINARY: PRT; 782 AA.  
 AC Q9NXK3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE CDNA FLJ20197 F1S, CLONE COLF0996.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COLON MUCOSA;  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
 RA Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T.,  
 RA Nakamura Y., Isono T., Sugano S.;  
 RT "NED0 human cDNA sequencing project."  
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK000204; BAA91007.1;  
 SO SEQUENCE 782 AA; 88616 MW; F472FE54F627CE8 CRC64;

Query Match 11.6%; Score 103.5; DB 4; Length 782;  
 Best Local Similarity 25.2%; Pred. No. 1.7;  
 Matches 40; Conservative 32; Mismatches 48; Indels 39; Gaps 7;  
 QY 54 RGLRAVME-----CRNTHLL---QQLTEAQGFQDVEAQ-AAATC-----NHTVMA 97  
 DB 284 RDSLHATAELQVROSLTHIALQEEELTRKVPDSLEPFRKQSLNMRREKVFA 343  
 QY 98 LMASLDAEKAQGGKVEELEGETTTLNKH-----LQDASAEVERLRRENOV 143  
 DB 344 LMVQLAKQLEHSDSVKOLKGVASLOEYVTSOSQEQAILQSLQDKAAEVEVERMGAKG 403  
 QY 144 LSVRIA---ADKKYPSSODSSSAAPOLLIVLGLSA 177  
 DB 404 LQLELSRAQEARRW---QQQTASAEQRLVYNAVSS 438

RESULT 8  
 Q9NXJ4 PRELIMINARY: PRT; 782 AA.  
 AC Q9NXJ4;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CDNA FLJ20210 F1S, CLONE COLF1787.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COLON MUCOSA;  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
 RA Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T.,  
 RA Nakamura Y., Isono T., Sugano S.;  
 RT "NED0 human cDNA sequencing project."  
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AK000217; BAA91016.1; -  
SQ SEQUENCE 782 AA; 88643 MW; A510C8BCFA8247B CRC64;

## Query Match

11.6%; Score 103.5; DB 4; Length 782;

Best Local Similarity 25.2%; Pred. No. 1.7;  
Matches 40; Conservative 32; Mismatches 48; Indels 39; Gaps 7;

QY 54 RDGLRAVME-----CRNVTHLL---QOELTEAOKGFODVEAO-AATC-----NHVYMA 97  
DB 284 RDSLHATAVLLQVAVGSLTHLALQEEELTRKQVPSDSLSEETFRKQCSLNRREKYFA 343  
QY 98 LMSLDEKAGOKKVEBELSEITTLNKK-----LQDASAEYERLRRENOV 143  
DB 344 LMVQLKAQLEHSDSVKOLKQVAVSLQEKYTSQSEDAIILQRSIQDKRAAEVEYRMGAKG 403  
QY 144 LSVRI-----ADKKYPPSSODSSAAPOLLIVILGLSA 177  
404 LQELSTRAGAEARROW-----QOQTASAEBOQLRVYNAVSS 438

## RESULT 9

028298

ID 028298 PRELIMINARY; PRT; 1534 AA.

AC Q28298;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE RIBOSOME RECEPTOR.  
GN P180.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95310363; PubMed=7790375;  
RA Wenker E.E., Sun Y., Savitz A.J., Meyer D.I.;  
RT "Functional characterization of the 180 kDa ribosome receptor in  
RT vivo."  
RL J. Cell Biol. 130:29-39(1995).  
DR EMBL: X87224; CA60676.1; -  
KW Receptor.  
SQ SEQUENCE 1534 AA; 164586 MW; B343BCF12656F3C5 CRC64;

## Query Match

11.5%; Score 102.5; DB 6; Length 1534;

Best Local Similarity 28.6%; Pred. No. 4.5;  
Matches 26; Conservative 22; Mismatches 42; Indels 1; Gaps 1;

QY 55 DGLRAVMECRNVTHLLQOELTEAOKGFODVEAOAATCNHTVMALMSLDEKAGOKKVE 114  
DB 847 DAIVAKSLREVNKEELAEAKAKAAAGKVKQQLVAREOETITAVQANIEASVREHVEVQ 906  
QY 115 ELEGETTLNKKLQDA-SAEYERLRRENOV 144  
DB 907 QLOGKIRFTLOEOLGPNQTLARLOENSIL 937

## RESULT 10

096200

ID 096200 PRELIMINARY; PRT; 1627 AA.

AC 096200;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE AXONEME-ASSOCIATED PROTEIN GASP-180.  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.  
NCBI\_TaxID=5741;  
GN [1]  
RP SEQUENCE FROM N.A.  
RA Elenddorf H.G., Rohrer S.C., La Vigne E.A., Nash T.E.;

RT "Novel Axoneme-Associated Proteins in Giardia lamblia."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF400249; AAK91740.1; -  
SQ SEQUENCE 1627 AA; 179270 MW; 021EED9763907DCC CRC64;

## Query Match

11.5%; Score 102.5; DB 5; Length 1627;

Best Local Similarity 27.9%; Pred. No. 4.8;  
Matches 34; Conservative 23; Mismatches 60; Indels 5; Gaps 3;

QY 46 IKANSEACROGLRAVMECRNVTHLLQOELTEAOKGFODVEAOAATCNHTVMALMSL-DA 104  
DB 1284 LRESAEALQDKLHLSRRAADGDLQKLYEQLEKDSGAKELVAREATTIDELQRLDR 1343  
QY 105 EKAGOK-KVEBELSEITTLNKKLQDASAEYERLRRENOVSVRIADKKYPPSSODSSA 163  
DB 1344 EBYDDKERRIADDELAVLNDGLKDDAEIAELREG---LEAOPATYTYPPESGEVGD 1400  
QY 164 AA 165  
DB 1401 AA 1402

## RESULT 11

091Y73

ID 091Y73 PRELIMINARY; PRT; 304 AA.

AC 091Y73;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE MACROPHAGE GALACTOSE N-ACETYL-GALACTOSAMINE SPECIFIC LECTIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC014811; AML14811.1; -  
KW Lectin.  
SQ SEQUENCE 304 AA; 34622 MW; 2272E1ADA2C0262A CRC64;

## Query Match

11.4%; Score 101.5; DB 11; Length 304;

Best Local Similarity 23.2%; Pred. No. 0.86;  
Matches 36; Conservative 38; Mismatches 48; Indels 33; Gaps 7;

QY 24 LGIGIIVLLIIVIGPLIIFTIKANSEACRD--GLRAVMECRNVTHLLQOELTEAOKGF 81  
DB 41 LGSLILLVVVSVIG-----SONSQLRDIDGTLRAIID--NTTSKIRAE-----F 83  
QY 82 ODVAQAATCNHTVMALMSLDEKAGQO-----KVEELEGRI-----TTLNKKLQDAS 131  
DB 84 QSLDSRADSEKGISLKVDEHDHROLOAGRDLSQVTSLESTLEKREOALKTDLTDLT 143  
QY 132 AEYERLRRENOVSVRIADKKYPPSSODSSAAP 166  
DB 144 DHVQQLRKDLKALTCQLANLK-----NNGSEVACCP 174

## RESULT 12

09YHD5

ID 09YHD5 PRELIMINARY; PRT; 826 AA.

AC 09YHD5;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE MYOSIN HEAVY CHAIN (FRAGMENT).  
HNC-4.  
OS Rana catesbeiana (Bull frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.

[illegible][illegible]

OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1].  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97205278; PubMed=9052782;  
RA Sawamura T., Kume N., Aoyama T., Moriaki H., Hoshikawa H., Alba Y.,  
Tanaka T., Miwa S., Katsura Y., Kita T., Masaki T.;  
RT "An endothelial receptor for oxidized low-density lipoprotein.";  
RL Nature 386:73-77(1997).  
DR EMBL: D89049; BAA19005.1; .  
DR HSSP: P20693; 1HLJ.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; CLECT; 1.  
DR PROSITE: PS50041; C\_TYPE\_LLECTIN\_2; 1.  
KM Receptor.  
SQ SEQUENCE 270 AA; 30892 MW; 6055B6881AD7053D CRC64;

Query Match 11.2%; Score 99.5; DB 6; Length 270;  
Best local Similarity 25.9%; Pred. No. 1.1;  
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;  
OY 23 LIGIGIIVLIIYILGVPLIITIKANSEACRDGLRAVMECRNVTLLQDELTEAQKGFQ 82  
DB 40 VICIGILVPIVILILQISQVSDLIKQO-----ANITH--QEDILEGQ---- 80  
OY 83 DVEAQAATCNHTWALMASIDAEKAGQOKVEELEGITTLNKHLDASAEVERLRRENO 142  
DB 81 -----IIAQRSEKS-AQESOKELKEMITFLAHKLDEKSKKLMELHRNL 124  
OY 143 VLSVRIADKKRY--PSSOD 159  
DB 125 NLQEVLEKANYSGPCPD 143

Search completed: June 4, 2002, 15:27:55  
Job time: 128 sec



6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529

QY 1 MASTSYDYCRVPMEDGDKRCKLLLGIGILVLLIVILGN

DB 1 MASTSDYCVPHEDDGRCKLLIGITLVLLIVGVLIIFTTKANSEACRDIKAV 60  
 QY 61 MECRNATHLLQOELTEAKQGFQDVEAQAATCNHTVMALMASLDAEKAGQKVEELEGEI 120  
 DB 61 MECRNATHLLQOELTEAKQGFQDVEAQAATCNHTVMALMASLDAEKAGQKVEELEGEI 120  
 QY 121 TLTNHLKQDASAEVERLRRENOYLSVRIADKKYTPSSQDSSAAAPOLLIVLGLSLAQ 180  
 DB 121 TLTNHLKQDASAEVERLRRENOYLSVRIADKKYTPSSQDSSAAAPOLLIVLGLSLAQ 180  
 RESULT 2  
 KUCR\_RAT STANDARD; PRT; 550 AA.  
 ID P10716;  
 AC 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DT Kupfer cell receptor.  
 KCLR.  
 Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-104.  
 RX MEDLINE=88227939; PubMed=2836387;  
 RA Hoyle G.W., Hill R.L.;  
 RT "Molecular cloning and sequencing of a cDNA for a carbohydrate  
 RT binding receptor unique to rat Kupfer cells.";  
 RL J. Biol. Chem. 263:7487-7492(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91107689; PubMed=1846367;  
 RA Hoyle G.W., Hill R.L.;  
 RT "Structure of the gene for a carbohydrate-binding receptor unique to  
 RT rat Kupfer cells.";  
 RL J. Biol. Chem. 266:1850-1857(1991).  
 CC -1- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCCOSE.  
 CC COULD BE INVOLVED IN ENDOCYTOSIS.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -1- TISSUE SPECIFICITY: KUPFER CELLS.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J03734; AAA1472.1; -  
 CC EMBL: M55532; AAA40892.1; -  
 CC PIR: A28166; A28166.  
 CC PIR: A38674; A38674.  
 CC HSSP: P20693; 1HLJ.  
 CC InterPro: IPR000017; Syntaxin.  
 CC InterPro: IPR001304; lectin.c.  
 CC Pfam: PF00059; lectin.c; 1.  
 CC SMART: SM00034; CLECT; 1.  
 CC SMART: SM00503; SYN; 1.  
 CC PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.  
 CC PROSITE: PS50041; C-TYPE\_LLECTIN\_2; 1.  
 CC Receptor: Transmembrane; Glycoprotein; Lectin; Signal-anchor;  
 CC Endocytosis.  
 CC  
 CC DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-TI MEMBRANE PROTEIN)  
 CC  
 CC DOMAIN 70 550  
 CC DOMAIN 438 538 EXTRACELLULAR (POTENTIAL).  
 CC DISULFID 440 536 C-TYPE LECTIN (SHORT FORM).  
 CC DISULFID 516 528 BY SIMILARITY.

FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 550 AA; 61104 MW; 9358A6CF4C306270 CRC64;  
 Query Match 11.5%; Score 102; DB 1; Length 550;  
 Best Local Similarity 28.0%; Pred. No. 0.57;  
 Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;  
 QY 46 IKANSEACRDLGRAVMECRNVTHLLQOELTEAKQGFQDVEAQAATCNHTVMALMASLDAE 105  
 DB 256 ISAEIQAMRDQMGRAG-----EMTSLKQDELTLTAQIQNNNGHLEQDTQIQGL 305  
 QY 106 KAQGQKVEELEGEITLTNHLKQDASAEVERLRRENOYLSVRIADKKYTPSSQDSSAA 165  
 DB 306 KAQ-LKSTSSLNSQIEVNGKLDKSSRELQTLRD---LSDVSALKSNVQMLQSLQAK 361  
 QY 166 POLIVLGLSLA 177  
 DB 362 AEVQSLKTGLEA 373  
 RESULT 3  
 KICR\_MOUSE STANDARD; PRT; 422 AA.  
 ID P05784; Q61766;  
 AC 01-NOV-1988 (Rel. 09, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Keratin, type I cytoskeletal 18 (Cytokeratin 18) (Cytokeratin endo B)  
 GN (Keratin D).  
 GN Krt18 OR KRT1-18 OR KERD.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89196920; PubMed=2467843;  
 RA Ichinose Y., Morita T., Zhang F., Srinahsonggram S., Tondella M.L.C.,  
 RA Matsumoto M., Nozaki M., Matsushiro A.;  
 RT "Nucleotide sequence and structure of the mouse cytokeratin endob  
 RT gene.";  
 RL Gene 70:85-95(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Teratocarcinoma;  
 RA Alonso A., Weber T., Jorcano J.L.;  
 RT "Cloning and characterization of keratin D, a murine endodermal  
 RT cytoskeletal protein induced during in vitro differentiation of F9  
 RT teratocarcinoma cells.";  
 RL Roux's Arch. Dev. Biol. 196:16-21(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86085876; PubMed=2416755;  
 RA Singer P.A., Trevor K., Oshima R.G.;  
 RT "Molecular cloning and characterization of the Endo B cytokeratin  
 RT expressed in preimplantation mouse embryos.";  
 RL J. Biol. Chem. 261:538-547(1986).  
 RN [4]  
 RP SEQUENCE OF 1-131 FROM N.A.  
 RX MEDLINE=88255838; PubMed=2454868;  
 RA Oshima R.G., Trevor K., Shevinsky L.H., Ryder O.A., Cecena G.;  
 RT "Identification of the gene coding for the Endo B murine cytokeratin  
 RT and its methylated, stable inactive state in mouse nonepithelial  
 RT cells.";  
 RL Genes Dev. 2:505-516(1988).  
 CC -1- SUBUNIT: HETEROPTRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.  
 CC KERATIN 18 ASSOCIATES WITH KERATIN 8.

CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND  
 CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)  
 CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).  
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M2832; AAA37552.1; -  
 DR EMBL: M36376; AAA39373.1; -  
 DR EMBL: M11686; AAA39390.1; -  
 DR EMBL: Y00217; CAA68365.1; -  
 DR PIR: A25621; A25621.  
 DR PIR: A28428; A28428.  
 DR PIR: J70406; J70406.  
 DR SWISS-2DPAGE: P05784; MOUSE.  
 DR MGI: 96692; Krt1-18.  
 DR InterPro: IPR001664; IF.  
 DR InterPro: IPR002957; Keratin\_I.  
 DR Pfam: PF00038; filament\_1.  
 DR PRINTS: PR01248; TYPE1KERATIN.  
 DR PROSITE: PS00226; IF: 1.  
 KW Intermediate filament; Coiled coil; Keratin; Glycoprotein;  
 KW Acetylation.  
 FT INT\_MET 0 0 BY SIMILARITY.  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 1 70 HEAD.  
 FT DOMAIN 1 71 ROD.  
 FT DOMAIN 380 422 TAIL.  
 FT DOMAIN 71 106 COIL 1A.  
 FT DOMAIN 107 124 COIL 1B.  
 FT DOMAIN 125 216 COIL 1C.  
 FT DOMAIN 217 240 LINKER 12.  
 FT DOMAIN 241 379 COIL 2.  
 FT SITE 263 263 STUTTER.  
 FT SITE 323 323 STUTTER.  
 FT CARBOHYD 30 30 O-LINKED (GLCNAC) (BY SIMILARITY).  
 FT CARBOHYD 31 31 O-LINKED (GLCNAC) (BY SIMILARITY).  
 FT CARBOHYD 49 49 O-LINKED (GLCNAC) (BY SIMILARITY).  
 FT CONFLICT 133 133 L -> F (IN REF. 1).  
 FT CONFLICT 243 243 D -> N (IN REF. 2).  
 FT CONFLICT 252 252 A -> G (IN REF. 2).  
 SO SEQUENCE 422 AA; 47373 MW; 4D5B0E9C7732F2F3 CRC64.

Query Match 11.4%; Score 101; DB 1; Length 422;  
 Best Local Similarity 27.4%; Pred. No. 0.51; Mismatches 56; Indels 50; Gaps 8;  
 Matches 51; Conservative 27; Mismatches 56; Indels 50; Gaps 8;

QY 44 FTIKANSE-ACRD-----GLRAVMECRNVTLL-----LQDELTEAKRG---FOD 83  
 DB 156 FRVKETELAMQSVESDIHGKRVVDDPNTRQLRETEALKKELLFMKNNHBEVGG 215  
 QY 84 VEAQATCNHY-----MALMASIDAE-KAQQKQVLEL-----GELT 121  
 DB 216 LEAQIASSGLVYVDAPKPSODLSKIMADIRAQYELAKNRELDKYSQIESTTVVT 275  
 QY 122 TLNHLQDASAEVRLRRNOVLAVRIADKKYPSODSS-----SAAAPOLLIVLL 173  
 DB 276 TKSATIRDAETTLTLEKRLQTLLEIDDSMKNNINLENSGLDVEARKYKAQMEQLNGVLL 335  
 QY 174 GLSALL 179  
 DB 336 HLESEL 341

ID MGI\_MOUSE STANDARD: PRT; 304 AA.  
 AC P49300;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage  
 DE galactose/N-acetylgalactosamine-specific lectin) (MGL).  
 GN MGL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HEN;  
 RX MEDLINE=92268032; PubMed=1587794;  
 RA Sato M., Kawakami K., Osawa T., Toyoshima S.;  
 RT "Molecular cloning and expression of cDNA encoding a galactose/N-  
 RT acetylgalactosamine-specific lectin on mouse tumoricidal  
 RT macrophages.";  
 RL J. Biochem. 111:331-336(1992).  
 RN [2]  
 RP SEQUENCE OF 102-120 AND 137-151.  
 RC STRAIN=C3H/HEN;  
 RX MEDLINE=89197865; PubMed=3241002;  
 RA Oda S., Sato M., Toyoshima S., Osawa T.;  
 RT "Purification and characterization of a lectin-like molecule specific  
 RT for galactose/N-acetyl-galactosamine from tumoricidal macrophages.";  
 RL J. Biochem. 104:600-605(1988).  
 CC -1- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE  
 CC UNITS. MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL  
 CC MACROPHAGES AND TUMOR CELLS.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -1- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED  
 CC MACROPHAGES.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: S36676; AB22171.1; -  
 DR HSSP: P06734; IKCE.  
 DR MGI: 96975; Mgl.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR SMART: SM00034; CLECT; 1.  
 DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.  
 DR PROSITE: PS00041; C-TYPE\_LLECTIN\_2; 1.  
 KW Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.  
 FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).  
 FT TRANSM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 57 304 (POTENTIAL).  
 FT DOMAIN 172 298 EXTRACELLULAR (POTENTIAL).  
 FT DISULFID 173 184 C-TYPE LECTIN (LONG FORM).  
 FT DISULFID 201 296 BY SIMILARITY.  
 FT DISULFID 274 288 BY SIMILARITY.  
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 304 AA; 34596 MW; 3F79CD12C34F5BCC CRC64.

Query Match 11.2%; Score 99.5; DB 1; Length 304;  
 Best Local Similarity 23.2%; Pred. No. 0.47; Mismatches 49; Indels 33; Gaps 7;  
 Matches 36; Conservative 37; Mismatches 49; Indels 33; Gaps 7;

QY 24 LGIGILVLLIIVLGVPLIIFTIKANSEACRD--GLRAVMECRNVTLLQDELTEAKRGF 81

Db 41 LGISLLLVVSVIG-----SONSLRDLGTLFATLD--NTTSKIAE-----F 83  
 Oy 82 ODEAQAATCNHTVYMLMAISDAEKAGQ-----KKVELEGRI-----TTLNKKIADAS 131  
 Db 84 QSDSRADSEKISISKVYEDHROBQLQGRDLSQKRVTSLESTVEKREBALATDLSDLT 143  
 Oy 132 AEVERLRRENOVLVRIADKKYPPSSODSSSAAP 166  
 Db 144 DHVOQLRKDLKALTCOLANLK-----NNGSEVACCP 174

## RESULT 5

VDP\_HUMAN STANDARD; PRT; 962 AA.  
 ID VDP\_HUMAN  
 AC 060763;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE General vesicular transport factor p115 (Transcytosis associated protein) (TAP) (Vesicle docking protein).  
 GN VDP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxId=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A., AND PHOSPHORYLATION SITE SER-942.  
 RX MEDLINE-98148093; PubMed-9478999;  
 RT "Phosphorylation of the vesicle docking protein p115 regulates its association with the Golgi membrane."  
 RL J. Biol. Chem. 273:5385-5388(1998).  
 CC -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCYSTERNAL TRANSPORT IN THE GOLGI STACK. IT IS REQUIRED FOR TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.  
 CC -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.  
 CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER; PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.  
 CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE; PHOSPHORYLATION PROMOTES DISSOCIATION.  
 CC -1- SIMILARITY: BELONGS TO THE VDP/USOL/YB1047C FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement and its not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D86326; BAA25300.1; -  
 CC MIM; 603344; -  
 DR InterPro: IPR000225; Armadillo.  
 DR PROSITE: PS50176; ARM\_REPEAT; UNKNOWN.1.  
 KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil; Phosphorylation.  
 FT DOMAIN 1 637 GLOBULAR HEAD.  
 FT 638 930 COILED COIL (POTENTIAL).  
 FT DOMAIN 935 962 ASP/GLU-RICH (ACIDIC).  
 FT MOD\_RES 942 942 PHOSPHORYLATION.  
 FT MUTAGEN 942 942 S->A: LOSS OF PHOSPHORYLATION.  
 SQ SEQUENCE 962 AA; 107906 MW; 2E748E2C1BCB942 CRC64;

Query Match 11.1%; Score 99; DB 1; Length 962;  
 Best Local Similarity 28.5%; Pred. No. 1.8;

Matches 35; Conservative 26; Mismatches 54; Indels 8; Gaps 3;  
 Oy 60 VMECRNVTHLLOOELPEAGQFQDVE-AQAATCNHTVYMLMAISDAEKAGQKKVELEG 118  
 Db 739 IEELKRNOELLQSLTEKDSMTENKSSQTSNGNESSALVSRDSE-----QVLELKO 792  
 Oy 119 EITTLNKKIADASAEVERLRRENO-VLSVRIADKKYPPSSODSSSAAPOLLIVLGLSA 177  
 Db 793 ELATLKSQLSQSVETTKLOTEKREKQELLQTEAFKASVEVGGETITLATTDTVEGRLSA 852  
 Oy 178 LLQ 180  
 Db 853 LLQ 855

## RESULT 6

MYSB\_CAEEL STANDARD; PRT; 1966 AA.  
 ID MYSB\_CAEEL  
 AC P02566;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain B (MHC B).  
 DE UNC-54 OR MYO-4.  
 GN Caenorhabditis elegans.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-83273600; PubMed-6576334;  
 RA Karn J., Brenner S., Barnett L.;  
 RT "Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain gene are not separated by introns."  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).  
 RN (2)  
 RP SEQUENCE OF 850-1966 FROM N.A.  
 RX MEDLINE-82272395; PubMed-7202124;  
 RA McLachlan A.D., Karn J.;  
 RT "Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle."  
 RL Nature 299:226-231(1982).  
 RN (3)  
 RP SEQUENCE OF 1876-1966 FROM N.A.  
 RX MEDLINE-83232892; PubMed-6571695;  
 RA Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S., Waterston R.H.;  
 RT "The genes sup-7 x and sup-5 III of C. elegans suppress amber nonsense mutations via altered transfer RNA."  
 RL Cell 33:575-583(1983).

-----  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).  
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C.ELEGANS.  
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: J01050; AAA28124.1; -  
 DR EMBL: V01494; CAA24738.1; -  
 DR PIR: A02992; MMKW.  
 DR HSSP: P08799; LMND.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00063; myosin\_head.1.  
 DR Pfam: PF02736; Myosin\_N.1.  
 DR Pfam: PF01576; Myosin\_tail.1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head.1.  
 DR SMART: SM00242; MYSC.1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM ATP-binding; Methylation; Alkylation; Multigene family.  
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.  
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).  
 FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).  
 FT DOMAIN 1165 1176 HINGE.  
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMM).  
 FT NP\_BIND 177 184 ATP (BY SIMILARITY).  
 FT DOMAIN 665 687 ACTIN-BINDING.  
 FT DOMAIN 769 783 ACTIN-BINDING.  
 FT MOD\_RES 128 128 METHYLATION (TR1-) (POTENTIAL).  
 FT MOD\_RES 705 705 ALKYLATION (SH-1).  
 FT MOD\_RES 715 715 ALKYLATION (SH-2).  
 FT CONFLICT 1337 1337 E -> R (IN REF. 2).  
 FT CONFLICT 1880 1880 I -> L (IN REF. 2).  
 FT SEQUENCE 1966 AA; 225125 MW; B6F0BB2FE27B67F CRC64;

Query Match 11.1%; Score 98.5; DB 1; Length 1966;  
 Best Local Similarity 24.7%; Pred. No. 4.2;  
 Matches 36; Conservative 21; Mismatches 56; Indels 31; Gaps 5;

QY 47 KANSEACRDLGRVMECRNVTHLLOELTEAOKFPDVEAOATCNHTVMALMASIDAEK 106  
 DB 1811 EEAFAALKGKRVIAKLEQVRLESELDGEORFODANKNGRADRVRELOFYDEDEK 1870  
 QY 107 AQ-----GOKKVELEGEITTLN-----HKQDASAEVERLRREN 141  
 DB 1871 KNEERLQDLIDKLOOKLKTKQKOVEAE-ELANLNLOKYKOLTHOLEDAERAD--QAEEN 1927  
 QY 142 QVLAVRI--ADKRYPPSSODSSAA 164  
 DB 1928 SLSKMRSKRASAVAPGLQSSASAA 1953

RESULT 7  
 MYH4\_RABIT STANDARD; PRT; 1938 AA.  
 AC 028641:  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last annotation update)  
 DE Myosin heavy chain, skeletal muscle, juvenile.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Skeletal muscle;  
 RA Maeda K., Hostinova E., Roesch-Kleinkauf A., Schuster H., Gasperik J.,  
 RT Wittinghofer A.;  
 RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit  
 RT skeletal muscle and a novel cosynthesis of S-1 fragment with the  
 RT essential and regulatory light chains.";

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATTER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: U32574; AAA74199.1; -  
 DR HSSP: P08799; LMND.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00063; myosin\_head.1.  
 DR Pfam: PF02736; Myosin\_N.1.  
 DR Pfam: PF01576; Myosin\_tail.1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head.1.  
 DR SMART: SM00015; IQ.1.  
 DR SMART: SM00242; MYSC.1.  
 DR PROSITE: PS50096; IQ.1.  
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KW Multigene family.  
 FT DOMAIN 1 783 MYOSIN HEAD-LIKE.  
 FT DOMAIN 784 813 IQ.  
 FT NP\_BIND 179 186 COILED COIL (POTENTIAL).  
 FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 774 774 ACTIN-BINDING (BY SIMILARITY).  
 FT MOD\_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).  
 FT MOD\_RES 130 130 METHYLATION (TR1-) (BY SIMILARITY).  
 FT MOD\_RES 552 552 METHYLATION (TR1-) (BY SIMILARITY).  
 FT MOD\_RES 756 756 METHYLATION (TR1-) (BY SIMILARITY).  
 FT MOD\_RES 698 698 ALKYLATION (SH-1) (BY SIMILARITY).  
 FT MOD\_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).  
 FT SEQUENCE 1938 AA; 223064 MW; D8A8A2EC5B182626 CRC64;

Query Match 11.0%; Score 98; DB 1; Length 1938;  
 Best Local Similarity 26.1%; Pred. No. 4.6;  
 Matches 37; Conservative 20; Mismatches 45; Indels 40; Gaps 5;

QY 47 KANSEACR-----DGLRVMECRNVTHLLOELTEAOKFPDVEAOATCNHTVMALM 99  
 DB 1366 KANSEVAQMTKYETAIOETLEAKKKLQRLDAD--EHEAVVNAKK----- 1414  
 QY 100 ASDLAERAGOKKVELEGEITTLN-----HKQDASAEVERLR 138  
 DB 1415 ASLEKTKRQNLNVEDLMIDVERTNAACAALDKKQNPDKILAEKMKHKEETAELEASQ 1474  
 QY 139 RENQVLSVRIAD-KKYPSSOD 159  
 DB 1475 KESRSLSSTEVFKYKNAYEESLD 1496

RESULT 8  
MYSD.CAEEL STANDARD: PRT: 1938 AA.  
ID MYSD.CAEEL 019674:  
AC P02567: 019674:  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain D (MHC D).  
GN MYO-1 OR P06C7.10.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Peloderae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2:  
RX MEDLINE=89178677; PubMed=2926820;  
Dibb N.J., Maruyama I.N., Krause M., Karn J.;  
"Sequence analysis of the complete Caenorhabditis elegans myosin  
heavy chain gene family.";  
J. Mol. Biol. 205:603-613(1989).  
RN [2]  
RP SEQUENCE OF 34-1795 FROM N.A.  
RX MEDLINE=83273600; PubMed=6576334;  
Karn J., Brenner S., Barnett L.;  
"Protein structural domains in the Caenorhabditis elegans unc-54  
myosin heavy chain gene are not separated by introns.";  
Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).  
RN [3]  
RP SEQUENCE OF 115-365 AND 1492-1763 FROM N.A.  
RX MEDLINE=85201409; PubMed=3888374;  
Karn J., Dibb N.J., Miller D.M.;  
"Cloning nematode myosin genes.";  
Cell Muscle Motil. 6:185-237(1985).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Gardner A., McMurray A.;  
Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE  
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
SUBFRAGMENT (S2).  
CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN  
C.ELEGANS.  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation-  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X08065; CAA30854.1; -  
DR EMBL: M37232; AAA28119.1; -  
DR EMBL: M37232; AAA28120.1; -  
DR EMBL: Z71261; CAA95848.1; -  
DR EMBL: Z71261; CAA95848.1; JOINED.  
DR EMBL: Z71261; CAA95806.1; -

DR EMBL: Z71261; CAA95806.1; JOINED.  
DR PIR: S02772; MKKW1.  
DR HSSP: P08799; 1MND.  
DR WormRep: R06C7.10; CE06253.  
DR InterPro: IPR004009; Myosin\_N.  
DR InterPro: IPR002928; Myosin\_Tail.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00063; myosin\_head.1.  
DR Pfam: PF02736; Myosin\_N.1.  
DR Pfam: PF01576; Myosin\_Tail.1.  
DR PRINTS: PRO0193; MYOSINHEAVY.  
DR ProDom: PD000355; myosin\_head.1.  
DR SMART: SM00242; MYSC.1.  
DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Multigene family.  
FT DOMAIN 1 845  
FT MYOSIN HEAD-LIKE  
FT DOMAIN 846 1938  
FT RODLIKE TAIL (S2 AND LMM DOMAINS).  
FT DOMAIN 846 1170  
FT ALPHA-HELICAL TAILPIECE (SHORT S2).  
FT DOMAIN 1171 1938  
FT LIGHT MEROMYOSIN (LMW).  
FT AN\_BIND 846 1938  
FT COILED COIL (POTENTIAL).  
FT DOMAIN 177 184  
FT ATP (BY SIMILARITY).  
FT DOMAIN 660 682  
FT ACTIN-BINDING.  
FT DOMAIN 764 778  
FT ACTIN-BINDING.  
FT MOD.RES 128 128  
FT METHYLATION (TR1-).  
FT MOD.RES 700 700  
FT ALKYLATION (SH-1).  
FT MOD.RES 710 710  
FT ALKYLATION (SH-2).  
FT CONFLICT 94 94  
FT F -> E (IN REF. 2).  
FT CONFLICT 98 98  
FT A -> R (IN REF. 2).  
FT CONFLICT 377 377  
FT V -> D (IN REF. 4).  
FT CONFLICT 389 390  
FT DV -> GD (IN REF. 2).  
FT CONFLICT 391 391  
FT W -> D (IN REF. 4).  
FT CONFLICT 408 408  
FT V -> N (IN REF. 2).  
FT CONFLICT 474 474  
FT L -> G (IN REF. 4).  
FT CONFLICT 577 577  
FT L -> F (IN REF. 4).  
FT CONFLICT 681 681  
FT I -> N (IN REF. 4).  
FT CONFLICT 1373 1373  
FT S -> D (IN REF. 4).  
FT CONFLICT 1659 1659  
FT E -> Q (IN REF. 3).  
SQ SEQUENCE 1938 AA; 223255 MW; 387399C8F634ACF4 CRC64;

Query Match 11.0%; Score 98; DB 1; Length 1938;  
Best Local Similarity 26.6%; Pred. No. 4.6;  
Matches 34; Conservative 24; Mismatches 58; Indels 12; Gaps 4;  
QY 51 EACRDGIRAVME---CRNVTHLLOQ--ELTEAO-KGFQDVEAOATCNHTVMAIMASLD 103  
DB 1324 KANDELMHEQDFHACKNLEHEDQCHELLEDINKKDIOQLSINSEIOWKARYE 1383  
QY 104 AEKAGQCKVVELEGETTINHLKQDASAEVERLRRENOVLSVRIADKKYTPSSQDSSA 163  
DB 1384 GEGIVGSEEELEIKRKQMNREVMDOELMSAA-----ONKVISEKAKGKLAEETEDARSD 1438  
QY 164 AAPQLLIV 171  
DB 1439 VDRHLTVI 1446

RESULT 9  
PU92.SCICO STANDARD: PRT: 286 AA.  
AC P22312;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE PU9F.11/9-2 protein precursor.  
GN 11/9-2.  
OS Sciatara coprophila (Fungus gnaf).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Sciaroidea;  
OC Sciaridae; Bradysia.  
OX NCBI\_TaxID=38358;  
RN [1]  
RP SEQUENCE FROM N.A.





DE Hypothetical 100.0 kDa protein M01A8.2 in chromosome III.  
 GN M01A8.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durdin R., Favello A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Shownkeen R.,  
 RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
 RA Wohldman P.;  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RT  
 RT  
 RL Nature 368:32-38(1994).  
 CC  
 CC -1- SIMILARITY: CONTAINS 1 CAP-GLY DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: 227081; CAA81607.1; -  
 CC PIR: S40998; S40998.  
 CC PIR: S40999; S40999.  
 CC WormPep: M01A8.2; GE03491.  
 CC InterPro: IPR000938; CAP-GLY.  
 CC DR Pfam: PF01302; CAP-GLY. 1.  
 CC DR PROSITE: PS00845; CAP-GLY 1; 1.  
 CC DR PROSITE: PS50245; CAP-GLY 2; 1.  
 CC KW Hypothetical protein: Coiled coil.  
 CC FT DOMAIN 39 81 CAP-GLY.  
 CC FT DOMAIN 522 696 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 729 756 COILED COIL (POTENTIAL).  
 CC SQ SEQUENCE 893 AA; 99997 MW; 464F2962B36C28B1 CRC64;  
 Query Match 10.9%; Score 96.5; DB 1; Length 893;  
 Best Local Similarity 24.8%; Pred. No. 2.6;  
 Matches 32; Conservative 29; Mismatches 39; Indels 29; Gaps 4;  
 OY 48 ANSEARQDGLRVAECCRNVTYHLLD--OELTEAOKGFQDV-----EQAQAATNHTYMALMAS 101  
 DB 566 SNOQVYRNHANAAY-ESLQKTHETQIAEKKEFEERNEARRAEACAMNRRHQKVAC 624  
 OY 102 LDAEKAQOGKQKVELE-----GEITTLNHLKIDASAVERLR 139  
 DB 625 LDKIEAEKQCEQDLNDRKVLQALANDCDHRNOMLTRKISSLOTALTEKMSAEKRLRO 684  
 OY 140 ENQVLSVRI 148  
 DB 685 KNQNLSTLOY 693  
 RESULT 12  
 P91\_SCICO STANDARD; PRT; 286 AA.  
 AC P22311;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE PUFF II/9-1 protein precursor.  
 GN II/9-1.  
 OS Sclera coprophila (Fungus gnaf).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Scleroidea;  
 OC Sciaridae; Brachyia.  
 OX NCBI\_TaxID=38358;  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=6980;  
 RX MEDLINE=90133907; PubMed=2614832;  
 RA Dibartolomeis S.M., Gerbl S.A.;  
 RA "Molecular characterization of DNA puf II/9A genes in Sclera  
 coprophila.";  
 RT J. Mol. Biol. 210:531-540(1989).  
 CC  
 CC -1- MISCELLANEOUS: THE DNA PUFF II/9 PROTEINS HAVE A PROPOSED  
 CC INTERMOLECULAR COILED COIL STRUCTURE WITH POSSIBLY INTERMOLECULAR  
 CC DISULFIDE BRIDGES FORMED BY NUMEROUS CYSTEINE RESIDUES IN POSITION  
 CC D OF THE HEPTAD REPEAT.  
 CC  
 CC -1- SIMILARITY: 76% IDENTICAL TO THE PUFF II/9-2 PROTEIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X51680; CAA35983.1; -  
 CC PIR: S07532; S07532.  
 CC DR Signal: Coiled coil; Glycoprotein.  
 CC KW Signal: Coiled coil; Glycoprotein.  
 CC FT SIGNAL 1 19 OR 21 (POTENTIAL).  
 CC FT CHAIN 20 286 PUFF II/9-1 PROTEIN.  
 CC FT DOMAIN 61 235 HELICAL (POTENTIAL).  
 CC FT CARBOHYD 156 156 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC SQ SEQUENCE 286 AA; 32034 MW; AA6A7B55F191BBID CRC64;  
 Query Match 10.8%; Score 96; DB 1; Length 286;  
 Best Local Similarity 30.0%; Pred. No. 0.84;  
 Matches 24; Conservative 16; Mismatches 36; Indels 4; Gaps 1;  
 OY 59 AWECCRNVTYHLLDQELTEAOKGFQDVEAQAATNHTYMALMASLDAEKAQOGKQKVELEG 118  
 DB 102 ALCECQNSLSLQTEITQELKLAQTQELANCKEA-----LANCKENARLKKIEELNC 157  
 OY 119 EITTLNHLKIDASAVERLR 138  
 DB 158 TITQLELEQCRARRERDQ 177  
 RESULT 13  
 MYH8\_HUMAN STANDARD; PRT; 1937 AA.  
 AC P13535; Q14910;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUN-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).  
 GN MYH8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN  
 RP  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RC MEDLINE=90323631; PubMed=2373371;  
 RA Karsch-Mizirachi I., Peghali R., Jettliand L.A.;  
 RT "Generation of a full-length human perinatal myosin heavy-chain-  
 RT encoding cDNA.";  
 RT Gene 89:289-294(1990).



[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Skeletal muscle;  
RX MEDLINE=95324556; PubMed=7601129;  
RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,  
RT Stedman H.H., Rubinstein N.A.;  
"Characterization of a human perinatal myosin heavy-chain  
transcript.";  
RL Eur. J. Biochem. 230:1001-1006(1995).  
[3]  
RP SEQUENCE OF 502-1937 FROM N.A.  
RC TISSUE-Skeletal muscle;  
RX MEDLINE=90235862; PubMed=1691980;  
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,  
RT Arnold H.H.;  
"Identification of three developmentally controlled isoforms of human  
myosin heavy chains.";  
RL Eur. J. Biochem. 189:55-65(1990).  
[4]  
RP SEQUENCE OF 860-1937 FROM N.A.  
RX MEDLINE=89234168; PubMed=2715179;  
RA Feghali R., Leinwand L.A.;  
"Molecular genetic characterization of a developmentally regulated  
human perinatal myosin heavy chain.";  
RT J. Cell Biol. 108:1791-1797(1989).  
[5]  
RP SEQUENCE OF 1-46 FROM N.A.  
RA Esser K., Tidhar A., Myszkowski M.;  
"Isolation and characterization of the human perinatal MHC promoter.";  
RL Submitted (May-1998) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PFM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
SUBFRAGMENT (S2).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL: M36769; AAC17185.1; -  
DR EMBL: Z38133; CA86293.1; -  
DR EMBL: X51592; CA35841.1; -  
DR EMBL: M35250; AA36546.1; -  
DR EMBL: AF067143; AAC21557.1; -  
DR PIR: A30220; A30220.  
DR HSSP: P08799; 1LVK.  
DR MIM: 160741; -  
DR InterPro: IPR000048; IQ.  
DR InterPro: IPR004009; Myosin\_N.  
DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IQ; 1.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF02736; myosin\_N; 1.  
DR Pfam: PF01576; myosin\_tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR ProDom: PD000355; myosin\_head; 1.

DR SMART: SM00015; IQ; 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IQ; 1.  
KW Myosin, muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Multigene family;  
KW Calmodulin-binding.  
FT DOMAIN 1 780  
FT MYOSIN HEAD-LIKE.  
FT IQ.  
FT COILED COIL (POTENTIAL).  
FT NP\_BIND 181 188  
FT ATP.  
FT DOMAIN 658 680  
FT ACTIN-BINDING.  
FT MOD\_RES 132 132  
FT METHYLATION (PRI-) (POTENTIAL).  
FT MOD\_RES 698 698  
FT ALKYLATION (SH-1) (POTENTIAL).  
FT MOD\_RES 708 708  
FT ALKYLATION (SH-2) (POTENTIAL).  
FT CONFLICT 15 15  
FT A -> R (IN REF. 2).  
FT CONFLICT 970 970  
FT E -> Q (IN REF. 1 AND 4).  
FT CONFLICT 1072 1072  
FT M -> N (IN REF. 3).  
FT CONFLICT 1247 1247  
FT N -> H (IN REF. 1 AND 4).  
FT CONFLICT 1251 1252  
FT MC -> DSG (IN REF. 3).  
FT CONFLICT 1261 1261  
FT E -> G (IN REF. 1 AND 4).  
FT CONFLICT 1297 1297  
FT K -> Q (IN REF. 1 AND 4).  
FT CONFLICT 1377 1378  
FT KY -> NT (IN REF. 3).  
FT CONFLICT 1504 1505  
FT EN -> AH (IN REF. 1 AND 4).  
FT CONFLICT 1847 1847  
FT E -> D (IN REF. 1 AND 4).  
FT CONFLICT 1914 1914  
FT D -> H (IN REF. 2).  
SO SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;  
  
Query Match 10.7%; Score 95; DB 1; Length 1937;  
Best Local Similarity 25.2%; Pred. No. 7.9;  
Matches 33; Conservative 21; Mismatches 59; Indels 18; Gaps 4;  
  
QY 47 KANSEACR-----DGLRAVMECRNVTHLLOELTEAKGQDYDAQATCNHTVMALM 99  
DB 1366 KANSEVAQWRKYETDAIRTEELAEAKKKLAQRLOEAEHEVAANVACASLEKTKORIQ 1425  
QY 100 -----ASLDAEKAQC-----OKKVEELGEIITLHMKLODASAEVERLRRENOVSVRIA 149  
DB 1426 NEVEDLMIDVERSNACALDKKRNFDKVSSEMKRYETQALEASKESSRLSTELF 1485  
QY 150 D-KKYPSSQD 159  
DB 1486 KVKKNYEESLD 1496  
  
RESULT 14  
MYHB\_HUMAN STANDARD: PRT: 1972 AA.  
ID MYHB\_HUMAN AC P35749; 000396; P78422; 094944;  
AC 01-JUN-1994 (Rel. 29, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain, smooth muscle isoform (SMHMC).  
GN MYH1 OR KIAA0866.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN 11  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99425270; PubMed=10493829;  
RA Lofthus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,  
RA Fuhmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,  
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,  
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
"Genome duplications and other features in 12 Mb of DNA sequence from  
human chromosome 16p and 16q.";  
RT Genomics 60:295-308(1999).  
RL [2]  
RP SEQUENCE OF 1-1266 FROM N.A.  
RC TISSUE-Brain;  
RX MEDLINE=99156230; PubMed=10048485;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 5:355-364(1998).  
RN [3]  
RP SEQUENCE OF 885-1972 FROM N.A.  
RX MEDLINE-93263189; PubMed-7684189;  
RA Matsuda R., Yoshida M.C., Furutani Y., Imamura S., Kanda N.,  
RA Yanagisawa M., Masaki T., Takao A.;  
RT "Human smooth muscle myosin heavy chain gene mapped to chromosome  
RT region 16q12.";  
RL Am. J. Med. Genet. 46:61-67(1993).  
RN [4]  
RP SEQUENCE OF 1093-1972 FROM N.A.  
RC TISSUE=Hippocampus;  
RA Okajima K.;  
RT Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- TISSUE SPECIFICITY: SMOOTH MUSCLE, EXPRESSED IN THE UMBILICAL  
CC ARTERY, BLADDER, ESOPHAGUS AND TRACHEA.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- DISEASE: A CHROMOSOMAL REARRANGEMENT, KNOWN AS PERICENTRIC  
CC INVERSION INV(16)(P13Q22), PRODUCES A FUSION PROTEIN THAT CONSISTS  
CC OF THE 165 N-TERMINAL RESIDUES OF CBF-BETA (PEP2) WITH THE TAIL  
CC REGION OF MYH11. THIS REARRANGEMENT IS ASSOCIATED WITH ACUTE  
CC MYELOID LEUKEMIA OF M4EO SUBTYPE.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBFRAGMENT (S2).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF001548; AAC31665.1; -;  
DR EMBL: U91333; AAC35212.1; -;  
DR EMBL: AB020673; BAA74889.1; -;  
DR EMBL: D10667; -; NOT\_ANNOTATED\_CDS.  
DR EMBL: X69282; CAA49154.1; -;  
DR HSSP: P08799; IMMN.  
DR MIM: 160745; -;  
DR InterPro: IPR000048; IQ.  
DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR002017; Spectrin.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IQ; 1.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF01576; myosin\_tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR ProDom: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IQ; 2.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IQ; 1.  
DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
KW Multigene family; Proto-oncogene; Chromosomal translocation.

FT DOMAIN 1 785 MYOSIN HEAD-LIKE.  
FT 786 IQ.  
FT DOMAIN 815 COILED COIL (POTENTIAL).  
FT 844 1934 CARBOXYL-TERMINAL.  
FT 1935 1972 ATP (POTENTIAL).  
FT NP\_BIND 178 185 ACTIN-BINDING (BY SIMILARITY).  
FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).  
FT 762 776 ACTIN-BINDING (BY SIMILARITY).  
FT MOD\_RES 129 129 METHYLATION (TR-1) (POTENTIAL).  
FT MOD\_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).  
FT 711 711 ALKYLATION (SH-2) (POTENTIAL).  
FT 889 889 EER -> NSE (IN REF. 3).  
FT 1263 1266 ELDS -> TLSF (IN REF. 2).  
FT 1558 1558 T -> S (IN REF. 3).  
FT 1610 1611 KQ -> NE (IN REF. 3).  
FT 1786 1786 A -> S (IN REF. 4).  
FT 1958 1958 T -> L (IN REF. 3).  
SQ SEQUENCE 1972 AA; 227338 MW; 67665B2AEC1277 CRC64;  
Query Match 10.7%; Score 95; DB 1; Length 1972;  
Best Local Similarity 24.5%; Pred. No. 8;  
Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;  
QY 53 CCGCLAVMCCRVNTHLLODELTEAOKGPDVDAQAATCNHTYMAIASID----- 103  
DB 1268 CSDGERARAEALNDKVKHQLONEVESVTGMLNEAGKAIKLAKDVASLSOQDTQELLQEE 1327  
QY 104 -AKAGGOKKVEELEGITFLNKLDDASAEVRLRENOVLVRIADKK 152  
DB 1328 TROKLAVSTKRLDLEERNSLDQDLDEMAKONLEHISTLNIQSDSK 1377  
RESULT 15  
MYHB\_RABIT STANDARD; PRT; 1972 AA.  
AC P35748;  
DT 01-JUN-1994 (rel. 29, Created)  
DT 01-FEB-1996 (rel. 33, Last sequence update)  
DT 16-OCT-2001 (rel. 40, Last annotation update)  
DE Myosin heavy chain, smooth muscle isoform (SMMHC).  
GN MYH11.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92073350; PubMed-1961735;  
RA Rabin P., Kelly C., Perlasany M.;  
RT "Characterization of a mammalian smooth muscle myosin heavy-chain  
RT gene: complete nucleotide and protein coding sequence and analysis of  
RT the 5' end of the gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:10676-10680(1991).  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE  
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBFRAGMENT (S2).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its



THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2002, 15:25:07 ; Search time 16.33 Seconds  
(without alignments)  
1059.160 Million cell updates/sec

Title: US-09-828-217-1

Perfect score: 889

Sequence: 1 MASTSYDYCRVPMEDGDKRC.....SSAAPQLLVLLGLSALLQ 180

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR-71:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	889	100.0	180	2 A56836	bone marrow stroma
2	111	12.5	778	2 T30430	hypothetical prote
3	108	12.1	423	2 T59463	keratin, type I, c
4	107.5	12.1	461	2 H84099	cell wall-binding
5	107	12.0	143	2 G83955	flagellar protein
6	102.5	11.5	1534	2 A56734	ribosome receptor,
7	102	11.5	550	2 A28166	Kupffer cell recep
8	99.5	11.2	304	2 JX0209	lectin, galactose/
9	99	11.1	415	2 S35760	myosin heavy chain
10	98.5	11.1	1963	1 MMKW	myosin heavy chain
11	98	11.0	959	2 A55913	transcytosis-assoc
12	98	11.0	1938	1 MMKW1	myosin heavy chain
13	98	11.0	1938	2 A59293	skeletal myosin he
14	97	10.9	286	2 S07533	puff II/9A-2 prote
15	96.5	10.9	597	2 S40998	hypothetical prote
16	96.5	10.9	893	2 G88551	protein M01A8.2 I
17	96	10.8	286	2 S07532	puff II/9-1 protei
18	95.5	10.7	173	2 S76705	hypothetical prote
19	95	10.7	1937	2 T38055	myosin heavy chain
20	95	10.7	1972	1 A41604	myosin heavy chain
21	94.5	10.6	1388	2 S74245	serine/threonine-s
22	94	10.6	676	2 S00084	myosin heavy chain
23	93.5	10.5	140	2 H64629	myosin heavy chain
24	93.5	10.5	1938	1 S06005	myosin alpha heavy
25	93.5	10.5	1939	2 T48175	myosin heavy chain
26	93	10.5	359	2 T50712	TOP AP - chicken
27	93	10.5	848	2 A44972	paramyosin - nemat
28	93	10.5	879	2 A48575	paramyosin - nemat
29	93	10.5	1938	2 J05421	smooth muscle myos

30	93	10.5	1972	2 J05420	smooth muscle myos
31	92.5	10.4	392	2 G95258	secreted 45 kd pro
32	92.5	10.4	392	2 B98124	genital stress pro
33	92.5	10.4	1509	1 A27224	myosin heavy chain
34	92.5	10.4	1938	2 T49464	alpha cardiac myos
35	92	10.3	746	2 T47237	myosin II heavy ch
36	92	10.3	866	2 S04027	paramyosin - Caeno
37	92	10.3	872	2 T19296	hypothetical prote
38	91.5	10.3	244	2 T16913	beta-myosin heavy
39	91.5	10.3	764	2 T51302	myosin heavy chain
40	91.5	10.3	1039	2 S18199	myosin heavy chain
41	91.5	10.3	1934	2 T48153	myosin heavy chain
42	91.5	10.3	1935	1 A37102	myosin beta heavy
43	91.5	10.3	1935	1 S06006	myosin beta heavy
44	91.5	10.3	1939	1 A46762	myosin alpha heavy
45	91.5	10.3	2442	2 T08621	centrosome associa

#### ALIGNMENTS

```
RESULT 1
A56836
bone marrow stromal cell surface protein BST-2 - human
C:Species: Homo sapiens (man)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 21-Jul-2000
C:Accession: A56836
R:Shikawa, J.; Kaisho, T.; Tomizawa, H.; Lee, B.O.; Kobune, Y.; Inazawa, J.; Oritani
A:Title: Molecular cloning and chromosomal mapping of a bone marrow stromal cell surf
A:Reference number: A56836; MUID:95331788
A:Accession: A56836
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-180 <ISH>
A:Cross-references: GB:D28137; NID:q457563; PIDN:BA05679.1; PID:g506861
C:Genetics:
A:Gene: GDB:BSR2
A:Cross-references: GDB:409946; OMIM:600534
A:Map position: 19p13.2-19p13.2
C:Keywords: transmembrane protein

Query Match          100.0%; Score 889; DB 2; Length 180;
Best local Similarity 100.0%; Pred. No. 2.6e-62;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCLLIGILVLLIIVITLGVPLITFTIKNSACRDLRAV 60
    |||
DB 1 MASTSYDYCRVPMEDGDKRCLLIGILVLLIIVITLGVPLITFTIKNSACRDLRAV 60
    |||

QY 61 MECRNVTLLDQELTEAOKGFQDVEAQAATCNHTVMALMASIDAKAOKKVELEGEI 120
    |||
DB 61 MECRNVTLLDQELTEAOKGFQDVEAQAATCNHTVMALMASIDAKAOKKVELEGEI 120
    |||

QY 121 TTLNKKLDASAEVRLRENOVLSVRADKKYRSSODSSSAAPOLLIVLLGLSALLQ 180
    |||
DB 121 TTLNKKLDASAEVRLRENOVLSVRADKKYRSSODSSSAAPOLLIVLLGLSALLQ 180
    |||

RESULT 2
T30430
hypothetical protein ORF82 - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30430
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavick, J.M.; R
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A:Reference number: Z20836; MUID:99124785
A:Accession: T30430
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
```

A:Residues: 1-778 <KU2>  
A:Cross-references: EMBL:AF081810; PIDN:AACT0268.1

Query Match 12.5%; Score 111; DB 2; Length 778;  
Best Local Similarity 32.5%; Pred. No. 0.48;  
Matches 38; Conservative 22; Mismatches 47; Indels 10; Gaps 4;

QY 46 IKNSEACRQGLRAVMECR-NVTNLLQDELTEAKGQFQDVEAQAATNCNHTVMAIAMSUDA 104  
DB 520 LKAKSELNRD-LQAKAEQAQANARLQAEIDSUKR-----AESADADLRNVAQLEAEAES 574

QY 105 EKAQOGKRVKEELGETITTLNHLQDASAVERLRRENOVLVRADKKYPSQDSS 161  
DB 575 GAADLQNRARLRAEASSGLTRRQESNAEVLALRROKEDLERRTAAS---AAQDVS 627

RESULT 3  
59463  
A:Title: Type I, cytoskeletal - mouse  
A:Alternate names: endo B cytokeratin; keratin D  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 10-Dec-1999

C:Accession: I59463; A25621; A28428; J10406  
R:Alonso, A.; Weber, T.; Jorcano, J.L.  
Roux's Arch. Dev. Biol. 196, 16-21, 1987  
A:Title: Cloning and characterization of keratin D, a murine endodermal cytoskeletal pro

A:Reference number: I59463  
A:Accession: I59463  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-423 <RES>

A:Cross-references: GB:M36376; NID:q198587; PIDN:AAA39373.1; PID:q293682  
R:Singer, P.A.; Trevor, K.; Oshima, R.G.  
J. Biol. Chem. 261, 538-547, 1986

A:Title: Molecular cloning and characterization of the endo B cytokeratin expressed in F  
A:Reference number: A25621; MUID:86085876  
A:Accession: A25621

A:Molecule type: mRNA  
A:Residues: 1-243; 'D', 245-252; 'A', 254-423 <SIN>  
A:Cross-references: GB:M1686; NID:q198620; PIDN:AAA39390.1; PID:q293685

R:oshima, R.G.; Trevor, K.; Shevinsky, L.H.; Ryder, O.A.; Cecena, G.  
Genes Dev. 2, 505-516, 1988  
A:Title: Identification of the gene coding for the endo B murine cytokeratin and its me

A:Reference number: A28428; MUID:88255838  
A:Accession: A28428  
A:Molecule type: DNA  
A:Residues: 1-1137 <OSH>

A:Cross-references: GB:Y00217; NID:q50842; PIDN:CAA68365.1; PID:q50843  
R:ichinose, Y.; Morita, T.; Zhang, F.; Srimahasongram, S.; Tondella, M.L.C.; Matsumoto,  
Gene 70, 85-95, 1988  
A:Title: Nucleotide sequence and structure of the mouse cytokeratin endob gene.

A:Reference number: J10406; MUID:89196920  
A:Accession: J10406  
A:Molecule type: DNA  
A:Residues: 1-1133; 'F', 135-243; 'D', 245-252; 'A', 254-423 <ICH>

A:Cross-references: GB:M2832; NID:q340757; PIDN:AAA3552.1; PID:q532610  
C:Genetics:  
A:Gene: endob; KERD  
A:Introns: 132/3; 160/2; 212/3; 267/3; 309/3; 384/2  
C:superfamily: cytoskeletal keratin

C:keywords: coiled coil; intermediate filament

Query Match 12.1%; Score 108; DB 2; Length 423;  
Best Local Similarity 28.0%; Pred. No. 0.44;  
Matches 52; Conservative 28; Mismatches 56; Indels 50; Gaps 8;

QY 44 FTIKANSE-ACRD-----GLRAVMECRNVTNHL-----LQDELTEAKG-----FQD 83  
DB 157 FRAVYETELAMRQSVESDINGLRKRVVDQNTIRLQLETEIALEKELLFMKKNHEEBOG 216

QY 84 VEAQAATCNHTV-----MALMASLDAE-KAOGOKVVELE-----GEIT 121

DB 217 LERQIASGLTVEVDAPKQSDLSKIMANIRAOYEALGQNRREELDKYWSQJESTTYYT 276

QY 122 TLNHLQDASAVERLRRENOVLVRADKKYPSQDSS-----SAAPOLLIVL 173  
DB 277 TYSAEIRDAETTLTLETRTLQTLLEIDLSKMNONTLENSLQDVEARKYAKOEQLNGVLL 336

QY 174 GLSALL 179  
DB 337 HLESEL 342

RESULT 4  
H84099  
cell wall-binding protein BH3600 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: H84099  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: H84099  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-461 <STO>

A:Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BAB07319.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3600

Query Match 12.1%; Score 107.5; DB 2; Length 461;  
Best Local Similarity 26.1%; Pred. No. 0.52;  
Matches 31; Conservative 28; Mismatches 55; Indels 5; Gaps 2;

QY 33 IYIVLGVPLIIFIRKANSEACRQGLRAVMECRNVTNHLQDELTEAKGQFQDVEAQAATCN 92  
DB 5 ISLVAAAGLTFISLFSQSIEDA-KANSLQWQISDVOKERQKQKTEAEV----- 59

QY 93 HTVMAIAMSUDAERKAOGOKVKEELGETITTLNHLQDASAVERLRRENOVLVRADK 151  
DB 60 KEVERKLGDTAIEHLEDEVEETSKIQEKREIEVQAIELEKQIELEERAE 118

RESULT 5  
G83955  
flagellar protein required for flagellar formation flil [imported] - Bacillus halodur

C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: G83955  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G83955  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-143 <STO>

A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB06166.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: flil

Query Match 12.0%; Score 107; DB 2; Length 143;  
Best Local Similarity 22.1%; Pred. No. 0.17;  
Matches 31; Conservative 34; Mismatches 49; Indels 26; Gaps 4;

QY 23 LIGIGIVLLIIVLGVPLIIFIRKANSEACRQGLRAVMEC-----RNVTHLQDEL 74  
DB 6 LVNIMITLIVLTVGAVLIFVNFVNEDEQREPTIDEIRIAQSYETERTITNLSNDF 65

Oy 75 TEA-----QKGFQDVEAQAATCNHTVMAIAMSIDAEGAOKVEELEGITTLNKK 126  
Db 66 VRARFLHVNDKNNALQEQKROFQVNNIIIRSLAGMDASQSLGAGCIEKLEQ----- 118  
Oy 127 LODASAEVERLRRENOVLVS 146  
Db 119 LOD---DINALMOEGSVYK 135

## RESULT 6

A56734  
ribosome receptor, 180k - dog  
C:Species: Canis lupus familiaris (dog)  
C>Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: A56734  
R:Manter, E.E.; Sun, Y.; Savitz, A.J.; Meyer, D.I.  
J: Cell Biol. 130, 29-39, 1995  
A:Title: Functional characterization of the 180-kD ribosome receptor in vivo.  
A:Reference number: A56734; MUID:95310363  
A:Accession: A56734  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1534 <MAN>  
A:Cross-references: GB:X87224; NID:9984113; PIDN:CNA60676.1; PID:9984114  
C:Keywords: endoplasmic reticulum; membrane protein; protein biosynthesis  
F:198-743/Region: 10-residue repeats (N-Q-G-K-K-A-E-G-A-P)

Query Match 11.5%; Score 102.5; DB 2; Length 1534;  
Best Local Similarity 28.6%; Pred. No. 4.4;  
Matches 26; Conservative 22; Mismatches 42; Indels 1; Gaps 1;

Oy 55 DGLRVMCRNVTHTLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSIDAEGAOKVE 114  
Db 847 DAAYAKSLREYNNKLAELAKAKAAAGKAKVKQLVARQDETAVOARIEASTREHYEQ 906  
Oy 115 ELEGETTLNKKLODA-SAEVERLRRENOVL 144  
Db 907 QLOGKIRFLQOLENGPNTQLARLOQENSIL 937

## RESULT 7

A28166  
Kupffer cell receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 05-Nov-1999  
C:Accession: A38674; A28166  
R:Hoyle, G.W.; Hill, R.L.  
J: Biol. Chem. 266, 1850-1857, 1991  
A:Title: Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer  
A:Reference number: A38674; MUID:91107689

A:Accession: A38674  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-550 <HO2>  
A:Cross-references: GB:M5532; NID:g203362; PIDN:AAA40892.1; PID:g203363  
R:Hoyle, G.W.; Hill, R.L.  
J: Biol. Chem. 263, 7487-7492, 1988  
A:Title: Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor  
A:Reference number: A28166; MUID:88227939  
A:Accession: A28166  
A:Molecule type: mRNA  
A:Residues: 1-550 <HOY>  
A:Cross-references: GB:U03734; NID:g205050; PIDN:AAA41472.1; PID:g205051  
C:Superfamily: C-type lectin homology  
C:Keywords: transmembrane protein  
F:412-536/Domain: C-type lectin homology <LCH>

Query Match 11.5%; Score 102; DB 2; Length 550;  
Best Local Similarity 28.0%; Pred. No. 1.7;  
Matches 37; Conservative 20; Mismatches 61; Indels 14; Gaps 3;

Oy 46 IKANSEACRDLGRAVMECRNVTHTLQOELTEAOKGFQDVEAQAATCNHTVMAIAMSIDA 105  
Db 256 ISAEIQARNDGMRGE-----EMTSIKKLETLTAQIONANGLIEOTDTQIOGL 305  
Oy 106 KAQGGKVEELEGITTLNKKLODA-SAEVERLRRENOVLVSRIADKKYPPSSODSSAAA 165  
Db 306 KAO-LKSTSLNSQLEEVNNGKLKSSRELQTLRRD---LSDVSAKSNVMOQLNSLQKAK 361  
Oy 166 POLLIIVLGLSA 177  
Db 362 AEVOSLKTGLEA 373

## RESULT 8

JX0209  
lectin, galactose/N-acetylglactosamine-specific - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: JX0209; PX0009  
R:Sato, M.; Kawakami, K.; Osawa, T.; Toyoshima, S.  
J: Biochem. 111, 331-336, 1992  
A:Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylglact  
A:Reference number: JX0209; MUID:92268032  
A:Accession: JX0209  
A:Molecule type: mRNA  
A:Residues: 1-304 <SAT>  
A:Cross-references: GB:S36676; NID:g249360; PIDN:AMB22171.1; PID:g249361  
R:Oda, S.; Sato, M.; Toyoshima, S.; Osawa, T.  
J: Biochem. 104, 600-605, 1988  
A:Title: Purification and characterization of a lectin-like molecule specific for gal  
A:Reference number: PX0009; MUID:89197865  
A:Accession: PX0009  
A:Molecule type: protein  
A:Residues: 102-120;137,X',139-151 <ODA>  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: glycoprotein; lectin; macrophage; transmembrane protein  
F:36-61/Domain: transmembrane #status predicted <TRA>  
F:173-296/Domain: C-type lectin homology <LCH>  
F:74,166/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 11.2%; Score 99.5; DB 2; Length 304;  
Best Local Similarity 23.2%; Pred. No. 1.4;  
Matches 36; Conservative 37; Mismatches 49; Indels 33; Gaps 7;

Oy 24 LGIGLIVLIIIVLGVPLIIFTIKANSEACRD--GLRAVMECRNVTHTLQOELTEAOKGF 81  
Db 41 LGISLLLVVSVIG-----SONSLRDLGTLRATLD--NTTSKIKAE-----F 83  
Oy 82 QDVEAQAATCNHTVMAIAMSIDAEGAOKG-----KYVELEGEI-----TTLNKKLODAS 131  
Db 84 QSLDRADSPFEKGISLVAVYEDHROELQAGRDISOKVTSLESTVEKREQALKTDLSDLT 143  
Oy 132 AEVERLRRENOVLVSRIADKKYPPSSODSSAAAP 166  
Db 144 DHVQQLRDLKALTLQGLANLKK-----NNSSEVACCP 174

## RESULT 9

S35760  
ferrA protein precursor - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 26-Aug-1999  
C:Accession: S35760; A42711  
R:Podbielski, A.  
submitted to the EMBL Data Library, November 1992  
A:Reference number: S35760  
A:Accession: S35760  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-415 <POD>  
A:Cross-references: EMBL:X69324; NID:g311759; PIDN:CNA49165.1; PID:g311760  
R:Haanes, E.J.; Heath, D.G.; Cleary, P.P.

```

A:Residues: 1873-1963 <MT3>
A:Cross-references: GB:V01494; GB:J01049; NID:g6783; PIDN:CAA24738.1; PID:g6784
C:Genetics:
A:Gene: unc-54; CESP:F1IC3.3
A:Map position: 1
A:Interons: 21/3; 64/3; 111/3; 264/1; 525/3; 951/2; 1747/3; 1819/3; 1894/3
A:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolyase; methylated amino acid; muscle
F:84-775/Domain: myosin motor domain homology <MOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:662-684/Region: actin binding #status predicted
F:766-780/Region: actin binding #status predicted
F:848-1963/Domain: coiled coil #status predicted <COI>
F:848-1162/Region: S2
F:1163-1963/Region: light meromyosin
F:125/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F:180/Binding site: ATP (Lys) #status predicted
F:702,712/Active site: Cys #status predicted

Query Match      11.1%; Score 98.5; DB 1; Length 1963;
Best Local Similarity 24.7%; Pred. No. 12;
Matches 36; Conservative 21; Mismatches 58; Indels 31; Gaps 5;

OY   47 KANEACRDGIRAVEECNNYTHLLQDELTEROKGFQDYEAQAATCNHTVALMLASDAEK 106
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   1808 EAFAALAKGGKKVKIAKLEQRVRELESELDEGDQRFPDANKNLIGRADRVRELQFOVEDK 1867

OY   107 AQ-----GCKVEELGETTTLN-----HKLDASAVERELREN 141
    : ||| | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db   1868 KNFERLODLIDKLQKLTQQKYQEAE-ELANLNLYKQLTHOLEDAERAD--QEN 1924

OY   142 QVLIVRI---ADKKRYPPSSODSSSA 164
    : | : | : | : | : | : | : | : | : | : | : | : |
Db   1925 SLTKMRSKSRASAVAPGLQSASAA 1950

RESULT 11 transcytosis-associated protein p115 - rat
        A55913 transcytosis-associated protein p115 (Norway rat)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 05-Nov-1999
C:Accession: A55913
R:Barroso, M.; Nelson, D.S.; Settl, E.
Proc. Natl. Acad. Sci. U.S.A. 92, 527-531, 1995
A>Title: Transcytosis-associated protein (RAP)/p115 is a general fusion factor requir
A:Reference number: A55913; MUID:95132633
A:Accession: A55913
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-959 <BAR>
A:Cross-references: GB:I5589; NID:g558474; PIDN:AAC52151.1; PID:g558475
C:Keywords: membrane fusion; membrane trafficking

Query Match      11.0%; Score 98; DB 2; Length 959;
Best Local Similarity 30.1%; Pred. No. 6.1;
Matches 37; Conservative 21; Mismatches 57; Indels 8; Gaps 3;

OY   60 VMECRNVTHLLQDELTEROKGFQDY-EAQATCTNHNTVALMASDAEKAQGCKVEELEEG 118
    : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   739 IEELRSHQVLLQSLQSLAEADVITENLRISGYSGMSEQALATCSPRDAE-----QVAETLKQ 792

OY   119 EITLHNHKLQDASAEVERELRENOVLVRIND-KKYPSPSODSSSAAPAPELLIYLGLISA 177
    : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   793 ELIALKSQLCSQSLEITRLQTEINRELQRAETLAQSVPGESSELVPAKTDDVERLSA 852

OY   178 LIQ 180
    |||
Db   853 LIQ 855

RESULT 12
```



MMKWI  
 myosin heavy chain D [similarity] - Caenorhabditis elegans  
 N:Alternate names: myosin heavy chain I  
 N:Contains: myosin ATPase (EC 3.6.1.32)  
 C:Species: Caenorhabditis elegans  
 C>Date: 28-Feb-1986 #sequence\_revision 19-May-2000 #text\_change 19-Jan-2001  
 C:Accession: T21193; T23973; S02772; A02993  
 R:McMurray, A.  
 submitted to the EMBL Data Library, April 1996  
 A:Reference number: Z19388  
 A:Accession: T21193  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1938 <M12>  
 A:Cross-references: EMBL:Z71261; PIDN:CAA95806.1; GSPDB:GN00019; CESP:R06C7.10  
 A:Experimental source: clone F21C3  
 R:Gardner, A.  
 submitted to the EMBL Data Library, April 1996  
 A:Reference number: Z19825  
 A:Accession: T23973  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1938 <M12>  
 A:Cross-references: EMBL:Z71266; PIDN:CAA95848.1; GSPDB:GN00019; CESP:R06C7.10  
 A:Experimental source: clone R06C7  
 R:Diib, N.J.; Maruyama, I.N.; Krause, M.; Karn, J.  
 J Mol. Biol. 205, 603-613, 1989  
 A:Title: Sequence analysis of the complete Caenorhabditis elegans myosin heavy chain gen  
 A:Reference number: S02771; MUID:89178677  
 A:Accession: S02772  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-376, 'V', 378-390, 'V', 392-576, 'L', 578-680, 'T', 682-1938 <DIB>  
 A:Cross-references: EMBL:X08065; NID:96785; PIDN:CAA30854.1; PID:96786  
 R:Karn, J.; Brenner, S.; Barnett, L.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983  
 A:Title: Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy ch  
 A:Reference number: A93958; MUID:83273600  
 A:Accession: A02993  
 A:Molecule type: DNA  
 A:Residues: 24-93, 'E', 95-97, 'R', 99-376, 'V', 378-388, 'GDV', 392-407, 'N', 409-473, 'G', 475-576  
 C:Genetics:  
 A:Gene: myo-1; CESP:R06C7.10  
 A:Map position: 1  
 A:Introns: 23/3; 114/3; 229/1; 264/1; 320/1; 857/3; 1745/3; 1814/1; 1892/3  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle cc  
 F:87-773/Domain: myosin motor domain homology <MMOT>  
 F:177-184/Region: nucleotide-binding motif A (P-loop)  
 F:660-682/Region: actin binding #status predicted  
 F:764-778/Region: actin binding #status predicted  
 F:846-1938/Domain: coiled coil #status predicted <COI>  
 F:846-1160/Region: S2  
 F:1161-1938/Region: light meromyosin  
 F:128/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted  
 F:183/Binding site: ATP (Lys) #status predicted  
 F:700/710/Active site: Cys #status predicted

Query Match 11.0%; Score 98; DB 1; Length 1938;  
 Best Local Similarity 26.6%; Pred. No. 13;  
 Matches 34; Conservative 24; Mismatches 58; Indels 12; Gaps 4;

OY 51 EACRQGLRAVME-----CRVTHLTLQO--ELTFEQ--KGFDVEQAQATCNHTVWALMASLD 103  
 DB 1324 KAAEELHERQEPHACKNLEHLLQCHLEEQINGKDDIQRLSRINSELSQWAKRYE 1383  
 OY 104 AAKAQOKKVEELEGITTLNKLQDASAEVERLRENOVLVSRIDKKRYPPSSODSSA 163  
 DB 1384 GEGIVGSEELBELKKRQMRVNDLQALSA-----QNKVISELKKAKGLLAETEDARSD 1438  
 OY 164 AAPOLLIV 171  
 |::|

DB 1439 VDRHLTVI 1446

RESULT 13  
 A59293  
 skeletal myosin heavy chain - domestic rabbit

C:Species: Oryctolagus cuniculus  
 C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000  
 C:Accession: A59293  
 R:Maeda, K.; Hostilova, E.; Roess, Kleinlauf, A.; Schuster, H.; Gasperik, J.; Wittingh  
 submitted to GenBank, July 1995  
 A:Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal  
 A:Reference number: A59293  
 A:Accession: A59293  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-1938 <MAE>  
 A:Cross-references: GB:U32574; NID:9940232; PIDN:AAA74199.1; PID:9940233  
 A:Experimental source: strain New Zealand White; cell type skeletal muscle fiber type  
 C:Genetics:  
 A:Gene: MHC  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 F:89-769/Domain: myosin motor domain homology <MMO>

Query Match 11.0%; Score 98; DB 2; Length 1938;  
 Best Local Similarity 26.1%; Pred. No. 13;  
 Matches 37; Conservative 20; Mismatches 45; Indels 40; Gaps 5;

OY 47 KANSEACR-----DGLRAVMECRNVTHLLQDELTEAQKGFQDYEAQATCNHTVWALM 99  
 DB 1366 KANSEVAQWRTRYEDDAIRTELEBAKKKLAQRLQDA---EHVEAVNAKC----- 1414  
 OY 100 ASLDEAKAQCKVELESEITTLN-----HKLDASAEVERLR 138  
 DB 1415 ASLEKQKRLQNEVEDLMDVERTNAACAALDKQRNPDKILAEWKHKYEETHAELEAQ 1474  
 OY 139 REMOVSRIAD-KKYPSSOD 159  
 DB 1475 KESRSLSTEVFKVNAYESLD 1496

RESULT 14  
 S07533

puff II/9A-2 protein precursor - fungus gnat (Sciara coprophila)  
 C:Species: Sciara coprophila  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Jun-2000  
 C:Accession: S07533  
 R:DiBartolomeis, S.M.; Gerbi, S.A.  
 J Mol. Biol. 210, 531-540, 1989  
 A:Title: Molecular characterization of DNA puff II/9A genes in Sciara coprophila.  
 A:Reference number: S07532; MUID:90133907  
 A:Accession: S07533  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-286 <DIB>  
 A:Cross-references: GB:X51679; NID:910113; PID:q1405812  
 C:Genetics:  
 A:Map position: II/9A  
 C:Keywords: coiled coil; glycoprotein  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-286/Product: puff II/9A protein #status predicted <MAP>  
 F:156/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.9%; Score 97; DB 2; Length 286;  
 Best Local Similarity 25.6%; Pred. No. 2.1;  
 Matches 31; Conservative 24; Mismatches 36; Indels 30; Gaps 5;

OY 46 IKANSEACRQGLRAVMECR--RVTHL-----LQDELTEAQKGFQDYEAQATCNHTVWAL 98  
 DB 89 LKREKARQAKERAKKCECKNTENIKETIEQLKKELEAEQAKLEKCKKEIADCK----- 142

```
OY / 99 MASIDAEACOGKVEEJEGEITTLNHNKLO-----DASAEVRLRE----NOVLAV 146
      | : : ||| | | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 143 -----KENAKLNNIEBELNCITITPQLQEKLECRGRERNDLQCQODECKRKNKINCNNELINC 197

OY 147 R 147
Db 198 R 198
```

RESULT 15

hypothetical protein M01A8.2 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
C:Accession: S40998  
R:Hawkins, T.; Thomas, K.  
submitted to the EMBL Data Library, October 1993  
A:Reference number: S40997  
A:Accession: S40998  
Status: preliminary  
Molecule type: DNA  
A:Residues: 1-597 <HAN>  
A:Cross-references: EMBL:227081  
C:Genetics:  
A:Introns: 40/3; 76/1; 125/1; 199/1; 240/1; 274/1; 292/3; 392/3; 429/2; 486/2; 521/1

	Query Match	10.9%;	Score 96.5;	DB 2,	length 597;	
	Best Local Similarity	24.8%;	Pred. No. 4.9;	Mismatches 32;	Conservative 29;	Gaps 4.
OY	ANSEACRDLGRAVMCECRVTHLLO--QELTEAKGFGDY----EQAATCNHTVALMAS	101				
Dd	270 SNGQIVRRHNAHV-ESLQKTHETQLAEKKKEERNFEEERARREAVCYCMNNRHQKVAC	328				
OY	102 LDAEAQAGOKKVEELE-----GETTLNHHKLQDASAEVERLRK	139				
Dd	329 LDEIKSEIEKQCEOLVDKVKVLQALANDCDHRNOMLTTEKISSLOTALEKSAEMKEMLRO	388				
OY	140 ENOVLSTRI	148				
Dd	389 KNQNLSTIQV	397				

Search completed: June 4, 2002, 15:27:21  
Job time: 134 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 4, 2002, 15:24:47 ; Search time 13.06 Seconds  
(without alignments)  
336.647 Million cell updates/sec

Title: US-09-828-217-1

Perfect score: 889  
Sequence: 1 MASTSYDRCVPMEDGDKRC.....SSAAPOLLIVLGLSALIQ 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/backfill1est.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889	100.0	180	2 US-08-624-650-1	Sequence 1, Appli
2	99.5	11.2	270	2 US-09-055-095-4	Sequence 4, Appli
3	99.5	11.2	270	2 US-08-809-494-2	Sequence 2, Appli
4	99.5	11.2	270	4 US-09-352-302-2	Sequence 2, Appli
5	99.5	11.2	273	2 US-08-809-494A-4	Sequence 4, Appli
6	99.5	11.2	273	4 US-09-352-302-4	Sequence 4, Appli
7	95	10.7	885	2 US-08-533-306A-4	Sequence 4, Appli
8	95	10.7	885	2 US-08-742-923A-4	Sequence 4, Appli
9	93.5	10.5	1886	4 US-08-938-105-3	Sequence 4, Appli
10	92.5	10.4	292	2 US-08-688-342-4	Sequence 4, Appli
11	92.5	10.4	292	2 US-09-113-788-4	Sequence 4, Appli
12	89.5	10.1	316	4 US-09-111-470-4	Sequence 4, Appli
13	89.5	10.1	1939	4 US-09-310-167A-1	Sequence 1, Appli
14	88	9.9	288	3 US-08-312-949-4	Sequence 4, Appli
15	88	9.9	288	3 US-08-446-201-4	Sequence 4, Appli
16	88	9.9	619	1 US-08-465-746-2	Sequence 2, Appli
17	88	9.9	619	1 US-08-214-164-2	Sequence 2, Appli
18	88	9.9	619	2 US-08-467-852A-3	Sequence 3, Appli
19	88	9.9	619	2 US-08-246-636-2	Sequence 3, Appli
20	88	9.9	619	2 US-08-247-491A-3	Sequence 3, Appli
21	88	9.9	619	2 US-08-319-795-2	Sequence 2, Appli
22	88	9.9	619	2 US-08-468-985-2	Sequence 2, Appli
23	88	9.9	619	3 US-08-312-949-2	Sequence 2, Appli
24	88	9.9	648	1 US-08-072-070-2	Sequence 2, Appli
25	88	9.9	648	1 US-08-469-434-2	Sequence 2, Appli
26	88	9.9	648	1 US-08-214-222-2	Sequence 2, Appli
27	88	9.9	648	2 US-08-467-852A-2	Sequence 2, Appli

28	88	9.9	648	2 US-08-468-718-2	Sequence 2, Appli
29	88	9.9	648	2 US-08-247-491A-2	Sequence 2, Appli
30	88	9.9	648	3 US-08-446-201-3	Sequence 2, Appli
31	88	9.9	695	1 US-08-127-499A-23	Sequence 23, Appli
32	88	9.9	695	1 US-08-482-847-23	Sequence 23, Appli
33	87.5	9.8	289	1 US-08-072-070-4	Sequence 4, Appli
34	87.5	9.8	289	1 US-08-469-434-4	Sequence 4, Appli
35	87.5	9.8	289	1 US-08-214-222-4	Sequence 4, Appli
36	87.5	9.8	289	2 US-08-467-852A-5	Sequence 5, Appli
37	87.5	9.8	289	2 US-08-468-718-4	Sequence 4, Appli
38	87.5	9.8	289	2 US-08-247-491A-5	Sequence 5, Appli
39	87.5	9.8	477	1 US-08-402-217A-3	Sequence 3, Appli
40	87.5	9.8	477	1 US-08-700-178-3	Sequence 3, Appli
41	87.5	9.8	477	3 US-08-995-654-3	Sequence 3, Appli
42	87.5	9.8	667	4 US-09-071-709-9	Sequence 9, Appli
43	85.5	9.6	459	4 US-09-071-709-1	Sequence 1, Appli
44	85	9.6	344	6 5210183-2	Patent No. 5210183
45	85	9.6	683	6 5210183-3	Patent No. 5210183

#### ALIGNMENTS

RESULT 1  
US-08-624-650-1  
Sequence 1, Application US/08624650  
Patent No. 5914252  
GENERAL INFORMATION:  
APPLICANT: HIRANO, TOSHIO  
APPLICANT: KAISHO, TSUNEYASU  
TITLE OF INVENTION: MEMBRANE PROTEIN POLYPEPTIDE HAVING  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS: PRE-B CELL GROWTH-SUPPORTING ABILITY AND A GENE THEREOF  
ADDRESSER: P.C.  
ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/624,650  
FILING DATE: 22-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/01732  
FILING DATE: 14-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-281622  
FILING DATE: 15-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 7625-001-0 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 180 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-624-650-1

Query Match 100.0%; Score 889; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 4.3e-88;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSDYCVPMEDGKCKLLGIGLIVLLIIVLGVPLIIFTRKANSEACRDGLRAV 60  
DB 1 MASTSDYCVPMEDGKCKLLGIGLIVLLIIVLGVPLIIFTRKANSEACRDGLRAV 60  
QY 61 MECRNVTLLQOELTEAOKGFQDVEAQAATCNHTVMALMSLDAEKAGOKVVEEGEI 120  
DB 61 MECRNVTLLQOELTEAOKGFQDVEAQAATCNHTVMALMSLDAEKAGOKVVEEGEI 120  
QY 121 TITNHLKODASAEVERLRNRQVLSVRIADKKYPPSSQDSSAAAPOLLIVLGLSALLQ 180  
DB 121 TITNHLKODASAEVERLRNRQVLSVRIADKKYPPSSQDSSAAAPOLLIVLGLSALLQ 180

RESULT 2  
US-09-055-095-4  
Sequence 4, Application US/09055095  
Patent No. 5945308  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Patterson, Chandra  
APPLICANT: Cotley, Neil C.  
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/055,095  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
FILING DATE:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0500 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 270 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1902982  
US-09-055-095-4

Query Match 11.2%; Score 99.5; DB 2; Length 270;  
Best Local Similarity 25.9%; Pred. No. 0.0068;  
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGLIVLLIIVLGVPLIIFTRKANSEACRDGLRAVMECRNVTLLQOELTEAOKGFQ 82  
DB 40 VLICGLIVLIVLLIILQISOVSDLIKQO-----ANITH--QEDILEGQ----- 80

QY 83 DVEAQAATCNHTVMALMSLDAEKAGOKVVEEGEITITNHLKODASAEVERLRNRQ 142  
DB 81 -----ILAQRSEKS--AQESQKELKEIETFLAHKLDKSKKLMELHQRNL 124  
QY 143 VLSVRIADKKY--PSSQD 159  
DB 125 NLOEVLKEANYSGPCPD 143

RESULT 3  
US-08-809-494A-2  
Sequence 2, Application US/08809494A  
Patent No. 5962260  
GENERAL INFORMATION:  
APPLICANT: Masaki, Tomoo  
APPLICANT: Sawamura, Tatsuya  
TITLE OF INVENTION: Modified Low-Density Lipoprotein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAlay Fisher Nissen Goldberg & Kiel  
STREET: 261 Madison Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,494A  
FILING DATE: 24-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-321705  
FILING DATE: 30-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-214206  
FILING DATE: 31-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldberg, Jules E  
REGISTRATION NUMBER: 24408  
REFERENCE/DOCKET NUMBER: JG-YV-4363PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 986-4090  
TELEFAX: 212 818-9479  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 270 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-809-494A-2

Query Match 11.2%; Score 99.5; DB 2; Length 270;  
Best Local Similarity 25.9%; Pred. No. 0.0068;  
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;

QY 23 LIGIGLIVLLIIVLGVPLIIFTRKANSEACRDGLRAVMECRNVTLLQOELTEAOKGFQ 82  
DB 40 VLICGLIVLIVLLIILQISOVSDLIKQO-----ANITH--QEDILEGQ----- 80  
QY 83 DVEAQAATCNHTVMALMSLDAEKAGOKVVEEGEITITNHLKODASAEVERLRNRQ 142  
DB 81 -----ILAQRSEKS--AQESQKELKEIETFLAHKLDKSKKLMELHQRNL 124  
QY 143 VLSVRIADKKY--PSSQD 159  
DB 125 NLOEVLKEANYSGPCPD 143

```

RESULT      4
US-09-352-302-2
; Sequence 2, Application US/09352302
Patent No. 6197937
GENERAL INFORMATION:
APPLICANT: Sawamura, Tatsuya
APPLICANT: Masaki, Tomoo
TITLE OF INVENTION: Modified Low-Density Lipoprotein
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAlulay Fisher Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/352,302
FILING DATE: 12-JUL-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-321705
FILING DATE: 30-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/DOCKET NUMBER: JG-YX-4363PCT/D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 986-4090
TELEFAX: 212 818-9479
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-352-302-2

Query Match          11.2%; Score 99.5; DB 4; Length 270;
Best local similarity 25.9%; Pred. No. 0.0068;
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5

OY    23 LIGIGILVLIIIVILGVPLLIIFTIRANSEACHDGRAWECRNVTLLLOQLTFAQGFQ 82
       :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    40 VLCGLGLTVTLIILLIQLSQVSIDLIKQQ-----ANITR--QEDILEGQ---- 80

OY    83 DVEADQAICNRTVMALMASLDAEKAGQGKVVLEEGETITLLHKKLDODASAEVRLRRNQ 142
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    81 -----TLAQRREKS-AQSOKELKEMETTLAKHKIDEKSKIMELHRNL 124

OY    143 VLSTRIADKKYY--PSSD 159
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    125 NLGEVLKEANYSGPCPD 143

RESULT      5
US-08-809-494A-4
; Sequence 4, Application US/08809494A
Patent No. 5962260
GENERAL INFORMATION:
APPLICANT: Sawamura, Tatsuya

```

```

APPLICANT: Masaki, Tomoo
TITLE OF INVENTION: Modified Low-Density Lipoprotein
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mcaulay Fisher Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,494A
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-321705
FILING DATE: 30-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/DOCKET NUMBER: JG-YY-4363PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 986-4090
TELEFAX: 212 818-9479
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-809-494A-4

Query Match      11.2% Score 99.5 DB 2 Length 273:
Best Local Similarity 25.9% Pred. No. 0.0069:
Matches 36: Conservative 25 Mismatches 41 Indels 37 Gaps 5

QY 23 LIGISILVLIVLIGVPLITFTIKANSEACRDGLRAVMESCHNNVTHLLQOELTEAQGFQ 82
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 43 VLCLGLLVTVILLIQLDSVSLIRKQQ-----ANITH-QEDLLLEG----- 83

QY 83 DVEQAATCNHTMALMASLDAEKAQGOKRYAEELGETTLNHLKIODASAEVERLRRENO 142
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 84 -----ILQARRSEKS-AQESQKELKEMIFETLHAKRIDEKSKIMELHRQL 127

QY 143 VLSVRIDAKRKY--PSSD 159
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 128 NLQEYLKEANYSGPCPD 146

RESULT        6
US-09-352-302-4
Sequence 4, Application US/09352302
Patent No. 6197937
GENERAL INFORMATION:
APPLICANT: Sawamura, Tatsuya
APPLICANT: Masaki, Tomoo
TITLE OF INVENTION: Modified Low-Density Lipoprotein
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mcaulay Fisher Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York

```

STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/352,302  
FILING DATE: 12-JUL-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-321705  
FILING DATE: 30-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-214206  
FILING DATE: 31-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldberg, Jules E  
REGISTRATION NUMBER: 24408  
REFERENCE/DOCKET NUMBER: JG-VY-4363PCT/D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 986-4090  
TELEFAX: 212 818-9479  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-352-302-4

Query Match 11.2%; Score 99.5; DB 4; Length 273;  
Best Local Similarity 25.9%; Pred. No. 0.0069;  
Matches 36; Conservative 25; Mismatches 41; Indels 37; Gaps 5;  
DB 23 LIGIGILVLLIIVLGVPLIIFIKANSEACRGDLRAVMECRVTHLLOQELTEAQKGFQ 82  
43 VLCIGILVYLLILLOLSQSDLIKQQ-----NATH--QEDILGQ---- 83  
OY 83 DVPAQATCNHTYVMAIMASIDAKAGQKVELEGEITTLNHLQDASAEVERLRRENO 142  
84 -----LLAQRRSEK-AQESQKELKEMITLAKLDEKSKMLMELHRONL 127  
DB 143 VLSVRIADKKY--PSSOD 159  
128 NLOEVILKEANYSQPCPD 146

RESULT 7  
US-08-533-306A-4  
Sequence 4, Application US/08533306A  
Patent No. 5837457  
GENERAL INFORMATION:  
APPLICANT: Liu, Pu  
APPLICANT: Collins, Francis S.  
APPLICANT: Siciliano, Michael J.  
APPLICANT: Claxton, David  
TITLE OF INVENTION: Markers for Detection of Chromosome 16  
TITLE OF INVENTION: Rearrangements  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: USA  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/533,306A  
FILING DATE: September 25, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Dean F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-00869COB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810) 641-1600  
TELEFAX: (810) 641-0270  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 885 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-533-306A-4

Query Match 10.7%; Score 95; DB 2; Length 885;  
Best Local Similarity 24.5%; Pred. No. 0.11;  
Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;  
DB 53 CRGRLRAVMECRVTHLLOQELTEAQKGFQDVPAQATCNHTYVMAIMASID----- 103  
215 CSQGERARALNDKVKHLQNEVESYGMNLNABGAKAKLAKDVASLSQDQDQELLQEE 274  
OY 104 -AEKAGQKRVLELEGEITTLNHLQDASAEVERLRRENOVLSVRIADKK 152  
DB 275 TRQKLNVSTKLROLLEERNLSIQDLDEMEAKONLEHISTLNITQSDSK 324

RESULT 8  
US-08-742-923A-4  
Sequence 4, Application US/08742923A  
Patent No. 5869611  
GENERAL INFORMATION:  
APPLICANT: Liu, Pu  
APPLICANT: Collins, Francis S.  
APPLICANT: Siciliano, Michael J.  
APPLICANT: Claxton, David  
TITLE OF INVENTION: Markers for Detection of Chromosome 16  
TITLE OF INVENTION: Rearrangements  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: P.O. Box 828  
CITY: Bloomfield Hills  
STATE: MI  
COUNTRY: USA  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/742,923A  
FILING DATE: No. 5869611ember 1, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Dean F.  
REGISTRATION NUMBER: 36683  
REFERENCE/DOCKET NUMBER: 2115-00869DVC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810) 641-1600  
TELEFAX: (810) 641-0270  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 885 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-742-923A-4

Query Match 10.7%; Score 95; DB 2; Length 885;  
Best Local Similarity 24.5%; Pred. No. 0.11;  
Matches 27; Conservative 18; Mismatches 55; Indels 10; Gaps 1;

53 CRDGRVWECRNTHLQOELTEKQGFQDVEAOATCNHTVMAASLD-----103

215 CSDGRVWECRNTHLQOELTEKQGFQDVEAOATCNHTVMAASLD-----103

104 -AEKQGGKVELEGETITLHKLODASAEVERLRRENOVLVSRADKK 152

275 TRKLNVTSTKLQLEERNSLQDQDEMEKONLEHISTINQLSDSK 324

## RESULT 9

US-08-938-105-3

Sequence 3, Application US/08938105

Patent No. 6353151

GENERAL INFORMATION:

APPLICANT: Leinwand, Leslie A.

APPLICANT: Vikstrom, Karen L.

TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln St., Suite 3500

CITY: Denver

STATE: CO

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/938.105

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Crook, Wamell M.

REGISTRATION NUMBER: 31,071

REFERENCE/DOCKET NUMBER: 3595-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1886 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-938-105-3

Db 1808 KNLVRLQDLVDRQL 1822

## RESULT 10

US-08-688-342-4

Sequence 4, Application US/08688342

Patent No. 5871964

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Cocks, Benjamin G.

APPLICANT: Goli, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/688.342

FILING DATE: Filed Herewith

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0095-1 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 292 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1235724

US-08-688-342-4

Query Match 10.4%; Score 92.5; DB 2; Length 292;  
Best Local Similarity 24.6%; Pred. No. 0.043;  
Matches 32; Conservative 28; Mismatches 39; Indels 31; Gaps 4;

20 CKLL--LGIGIVLLIVIGVPLIFITKANSEACRDGIRAVWECRNTHLQOELTEA 77

39 CHLLSLGIGILLVIVICVGF-----QNKRFQDLVTLTATDSNFTSNVAEL--- 87

78 QKGFQDVEAOATCNHTVMAASLDKRAQGO-----KVEELEGETITL 123

88 ---QALTSQSSLEETIASLKAIEGFGKQERQAVHSEMLLRVQOLVODLKLTCOVATL 143

124 NHTLQDASAE 133

144 NNNGEASTE 153

## RESULT 11

US-09-113-788-4

Sequence 4, Application US/09113788

Patent No. 5969104

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

```

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0695
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-4

Query Match      10.1% Score 89.5; DB 4; Length 316;
Best Local Similarity 27.8%; Pred. No. 0.1;
Matches 30; Conservative 20; Mismatches 41; Indels 17; Gaps 3

QY    20 CKL-LGIGILVLLIYLGVPLIIITIKANSEACRGLRAVMRCRNVTHTLLOELTEA 77
      ||| |||||:|||::| | | | | | | | | | | | | | | | | | | | | | | |
Db    39 CHLLSLDGLILLIVICVGF-----QNSKRQRDLVLTLPDFSNSTSVTAIEI--- 87

QY    78 QKGFDVAQAATCNHTYMAALMASIDAERAGSQCKVELEGCTITLNH 125
      | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db    88 ---QALTSGSSLETETIASLKAEEGFKEROAGVSELQETHTKAH 131

RESULT 13
Sequence 1, Application US/09310187A
Patent No. 6358751
GENERAL INFORMATION:
APPLICANT: Benichou, Gilles
APPLICANT: Fedoseyeva, Eugenia
TITLE OF INVENTION: Involvement of Autoantigens in Cardiac
TITLE OR INVENTION: Craft Rejection
FILE REFERENCE: UCSF-090
CURRENT APPLICATION NUMBER: US/09/310,187A
CURRENT FILING DATE: 1999-05-12
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1939
TYPE: PRT
ORGANISM: Homo sapiens
US-09-310-187A-1

Query Match      10.1% Score 89.5; DB 4; Length 1939;
Best Local Similarity 24.4%; Pred. No. 1.3;
Matches 33; Conservative 28; Mismatches 51; Indels 23; Gaps 5
```



[illegible]

RESULT 14  
US-08-312-

```

Sequence 4, Application US/08312949
Patent No. 6027734

GENERAL INFORMATION:
APPLICANT: Billes, David E.
APPLICANT: Wu, Hong-Yin
TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
TITLE OF INVENTION: PNEUMOCOCCAL ANTIGENS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtiss, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,949
FILING DATE: 30-SEP-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2049
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

```

Query Match	9.9%;	Score 88;	DB 3;	Length 288;
Best Local Similarity	30.9%;	Pred. No. 0.13;		
Matches 34;	Conservative 15;	Mismatches 29;	Indels 32;	Gaps 5;

OY	60	VMECRNTHLLOOELTE-----AKGROVEAOATCNHTVYMALMSLAERKOGQ	110
		: :	
Dd	178	IABENOVNHRDGLKEIDESEDDYAKEGR-----APLOSKIDAKRKLKS	224
OY	111	K-----KVBLEGEITTLNNHKIODA--SAEVEVRRENOVLSYRIADRK	152
		: :	
Dd	225	KLEBSKDIDELLAIAIKLEOLILAAENNNDVEDUEFEG--LEKTIAARK	272

## RESULT 1.5

US-08-446-201-4  
; Sequence 4, Application US/08446201B

```

: Patent No. 6042838
:
: GENERAL INFORMATION:
:
: APPLICANT: BRILES, David E.
: APPLICANT: WU, Hong-Yin
: TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS FOR MUCOSAL ADMINISTRATION OF
: TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN A (PSPA)
: FILE REFERENCE: 454312-2018
: CURRENT APPLICATION NUMBER: US/08/446, 201B
: CURRENT FILING DATE: 1995-05-19
: EARLIER APPLICATION NUMBER: 08/312,949
: EARLIER FILING DATE: 1994-09-30
: EARLIER APPLICATION NUMBER: 08/246,636
: EARLIER FILING DATE: 1994-05-20
: EARLIER APPLICATION NUMBER: 08/048,896
: EARLIER FILING DATE: 1993-04-20
: EARLIER APPLICATION NUMBER: 07/835,658
: EARLIER FILING DATE: 1992-02-12
: EARLIER APPLICATION NUMBER: 07/656,773
: EARLIER FILING DATE: 1991-02-15
: NUMBER OF SEQ. ID NOS.: 4
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ. ID NO. 4
:
: LENGTH: 288
:
: TYPE: PRS
:
: ORGANISM: Streptococcus pneumoniae
:
: US-08-446-201-4

```

Query Match	9.9%	Score 88;	DB 3;	Length 288;
Best Local Similarity	30.9%;	Pred. NO. 0.13;		
Matches 34;	Conservative 15;	Mismatches 29;	Indels 32;	Gaps 5;

```

Qy      60  VMEERNATLHLOQJLSTE-----AKOGDVEAATACNHTVMALMSLAEKROGO 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      178  IAELENGVHRLGEKELDESESDYAKKEGR-----APLOGLKDAKKKLS 224

Qy      111  K-----KVELEBGEIITLHNKLDA--SAEVLKRENOVLSYRIDKK 152
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      225  KLEELSDKIDELDAEIKLEQLKAAAEENNVNDEYFEG--LEKTIYAKK 272

```

Search completed: June 4, 2002, 15:26:57  
Job time: 130 sec



[illegible]

XX A membrane polypeptide has been isolated from rheumatoid arthritis  
 CC patients. The polypeptide supports pre-B-cell growth and is useful  
 CC as an antigen for generating monoclonal antibodies for diagnosing  
 CC rheumatoid arthritis.

XX Sequence 180 AA;

Query Match 100.0%; Score 889; DB 16; Length 180;

Best Local Similarity 100.0%; Pred. No. 6, 2e-78;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDCRVPMEDGDKRCKLLGIGILVLLITVILGVPFLITFTKANSEACDGLRAY 60  
 Db 1 mastsydcrvpmedgdkrcklllgilglvllitvllgvpflifftkanseacdglray 60  
 QY 61 MEGCRNTHLLQOELTERAKGFQDVEAQAATCNFTVAMALSLDAEKAGQKVEELGEI 120  
 Db 61 mecrnthlllqgelteakgfgdveaqaatcnhtvamaalsdaekagqkveelegei 120  
 QY 121 TTNLHKLODASAVERLRRENQVLSVRIADKRYPPSSQDSSAAAPQLLIVLIGLSALLQ 180  
 Db 121 ttnlhkldasaeverlrrenqvlsvriadkkyppssqdsaaapqllivlignlsallq 180

# RESULT 2

AAW65771 ID AAW65771 standard; Protein; 180 AA.

XX AAW65771;

DT 20-NOV-1998 (first entry)

DE Protein recognised by anti-human HMI.24 antibody.

XX Cytotoxic antibody; anti-human HMI.24; lymphocytic tumours;

KW pre-B lymphoma; Burkitt's lymphoma; T-cell tumour; B-cell tumour.

XX Homo sapiens.

PN MO9835698-A1.

PD 20-AUG-1998.

PF 12-FEB-1998; 98WO-JP00568.

PF 12-FEB-1997; 97JP-0041410.

PA (CHUS ) CHUGAI SEIYAKU KK.

PI Koishihara Y, Yoshimura Y;

DR WPI: 1998-456869/39.

DR N-PSDB; AAV07579.

PT Treatment of lymphocytic tumours using cytotoxic antibody - binding  
 to specific antigen such as HMI.24 and effective against T-cell  
 PT tumours and B-cell tumours other than myeloma

XX Claim 1: Page 44-45; 82pp; Japanese.

CC The protein having the amino acid sequence below is bound specifically  
 CC by a cytotoxic antibody which can be used in the treatment of lymphocytic  
 CC tumours, including T-cell tumours and B-cell tumours other than myeloma.  
 CC The antibody is preferably monoclonal and has ADCC or CDC type  
 CC cytotoxicity. It may be chimeric or humanised, and preferably contains  
 CC a human antibody constant region C gamma (such as C gamma 1 or C  
 CC gamma 3). A preferred antibody is an anti-human HMI.24 antibody or an  
 CC antibody which binds to an epitope recognising anti-human HMI.24  
 CC lymphocytic tumours such as acute or chronic B lymphocytic leukaemia,  
 CC pre-B lymphoma, Burkitt's lymphoma, or acute or chronic T lymphocytic

CC Leukemia.

XX Sequence 180 AA;

Query Match 100.0%; Score 889; DB 19; Length 180;

Best Local Similarity 100.0%; Pred. No. 6, 2e-78;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDCRVPMEDGDKRCKLLGIGILVLLITVILGVPFLITFTKANSEACDGLRAY 60  
 Db 1 mastsydcrvpmedgdkrcklllgilglvllitvllgvpflifftkanseacdglray 60  
 QY 61 MEGCRNTHLLQOELTERAKGFQDVEAQAATCNFTVAMALSLDAEKAGQKVEELGEI 120  
 Db 61 mecrnthlllqgelteakgfgdveaqaatcnhtvamaalsdaekagqkveelegei 120  
 QY 121 TTNLHKLODASAVERLRRENQVLSVRIADKRYPPSSQDSSAAAPQLLIVLIGLSALLQ 180  
 Db 121 ttnlhkldasaeverlrrenqvlsvriadkkyppssqdsaaapqllivlignlsallq 180

# RESULT 3

AAW62207 ID AAW62207 standard; Protein; 180 AA.

XX AAW62207;

DT 21-SEP-1998 (first entry)

DE Humanised anti-HMI.24 antibody polypeptide.

XX Mouse: human; humanised; anti-HMI.24 antibody; myeloma; FR; CDR;

KW Framework region; complementarity determining region; antigenicity.

XX Synthetic.

OS Mus sp.

PN MO9814580-A1.

PD 09-APR-1998.

PF 03-OCT-1997; 97WO-JP03553.

PF 04-OCT-1996; 96JP-0264756.

PA (CHUS ) CHUGAI SEIYAKU KK.

PI Koishihara Y, Kosaka M, Ohtomo T, Ono K, Tsuchiya M;

DR WPI: 1998-286421/25.

DR N-PSDB; AAV39359.

PT Humanised anti-HMI.24 antibody - for treatment of myeloma

XX Claim 81: Page 151-153; 210pp; Japanese.

CC A humanised anti-HMI.24 antibody has been developed which comprises  
 CC human L and H chain C regions, and L and/or H chain V regions  
 CC containing material originating in mouse anti-HMI.24 antibody. The V  
 CC regions contain framework (FR) regions of human origin and  
 CC complementarity determining regions (CDR) of mouse origin, leading to  
 CC a reshaped humanised antibody. The C regions are human Ck (L-chain) and  
 CC human C gamma (especially C gamma 1) (H-chain). The FR regions of the  
 CC L chain V region are derived from human subtype HSG1 (e.g. from human  
 CC antibody R61) and the FR regions of the H chain V region are derived  
 CC from human subtype HSG1 (e.g. FR1-3 from human antibody H63 and FR4  
 CC from human antibody JH6). The present sequence represents an antibody  
 CC polypeptide from the present invention. The antibodies are used for  
 CC the treatment of myeloma, especially by injection, intravenously,  
 CC intramuscularly or subcutaneously. The antibodies are used at 0.01-1000

CC (especially 5-100) mg/kg body weight. The humanised antibody has low  
 CC antigenicity and is therefore effective therapeutically in humans.  
 XX  
 SQ Sequence 180 AA:

Query Match 100.0%; Score 889; DB 19; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYGKRVPMEDGDKRCKLLGIGILVLLIIVILGVPILFTTKANSEACRDLRAV 60  
 DB 1 mastsydygcrvpmedgdkrcklllgigilvllilvllgvpilfttkanseacrdglrav 60  
 QY 61 MECRANTHLLQOELTFPAOKGFQDVEAOAATCNHTVMAIMASIDAEKAGQKKVEELEGET 120  
 DB 61 mecrantvhlhllqeltfpaqgfyfgdveaqaatcnhtvmaimasidaekagqkkveeleget 120  
 QY 121 TTLNHLKODASAEVERLRRENOVLSVRADKKRYPPSSODSSSAAPOLLIVLGLSALLQ 180  
 DB 121 tllnhklqdasaeverlrrengvlsvradiakkyppssqsssaapqllilvllglisallq 180

## RESULT 4

AAV33202  
 ID AAV33202 standard; Protein; 180 AA.

AC AAV33202;  
 DT 22-NOV-1999 (first entry)  
 XX  
 DE Human HML.24 antigenic protein.

KW Antigenic protein; HML.24; splice variant; promoter; antirheumatic;  
 KW antiarthritis; bone marrow; tumour cell; drug development; treatment;  
 KW myeloma; rheumatoid arthritis; human.

OS Homo sapiens.

PN WO9943803-A1.

PD 02-SEP-1999.

PF 25-FEB-1999; 99WO-JP00884.

PR 25-FEB-1998; 98JP-0060617.

PR 24-MAR-1998; 98JP-0093883.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Ohtomo T, Tsuchiya M, Koishihara Y, Kosaka M;

DR WPI; 1999-550869/46.

DR N-PSDB; AA209726.

XX Genomic DNA encoding HML.24 antigen protein as well as splicing  
 PT variants, useful e.g. in development of drugs for treating myeloma and  
 PT rheumatoid arthritis

PS Example 1; Fig 1-2; 83pp; Japanese.

CC This invention describes a novel human antigenic protein, HML.24,  
 CC its encoding nucleic acid, splice variants and promoter region. The  
 CC products of the invention have antirheumatic and antiarthritis activity.

CC The DNA of the invention is isolated from bone marrow tumour cells,  
 CC which can be used to study the expression of HML.24 antigen, promoter  
 CC activity of its promoter region, and in development of drugs in treating

CC e.g. myeloma and rheumatoid arthritis. This sequence represents the  
 CC human HML.24 antigenic protein described in the invention.

XX  
 SQ Sequence 180 AA:

Query Match 100.0%; Score 889; DB 20; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYGKRVPMEDGDKRCKLLGIGILVLLIIVILGVPILFTTKANSEACRDLRAV 60  
 DB 1 mastsydygcrvpmedgdkrcklllgigilvllilvllgvpilfttkanseacrdglrav 60  
 QY 61 MECRANTHLLQOELTFPAOKGFQDVEAOAATCNHTVMAIMASIDAEKAGQKKVEELEGET 120  
 DB 61 mecrantvhlhllqeltfpaqgfyfgdveaqaatcnhtvmaimasidaekagqkkveeleget 120  
 QY 121 TTLNHLKODASAEVERLRRENOVLSVRADKKRYPPSSODSSSAAPOLLIVLGLSALLQ 180  
 DB 121 tllnhklqdasaeverlrrengvlsvradiakkyppssqsssaapqllilvllglisallq 180

## RESULT 5

AAV32765  
 ID AAV32765 standard; Protein; 180 AA.

AC AAV32765;  
 DT 25-OCT-1999 (first entry)  
 XX  
 DE Soluble HML.24 antigenic protein sequence.

KW HML.24; antigenic protein; antibody; immunoassay; diagnosis; cancer;  
 KW immune disorder; immunotherapy monitoring.

OS Homo sapiens.

PN WO9943703-A1.

PD 02-SEP-1999.

PF 25-FEB-1999; 99WO-JP00885.

PR 25-FEB-1998; 98JP-0060613.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Koishihara Y, Ozaki Y;

DR WPI; 1999-518836/43.

DR N-PSDB; AA210917.

PT Immunoassay of anti-HML.24 antibody or soluble HML.24 antigen,  
 PT useful for diagnosis of immune disorders and cancer

PS Example 3; Fig 14-15; 138pp; Japanese.

CC This sequence represents a human soluble HML.24 antigenic protein. The  
 CC invention relates to an immunochemical assay of anti-HML.24 antibody by  
 CC use of a soluble HML.24 antigenic protein, or an immunochemical assay of  
 CC the soluble antigen by use of the antibody. The immunoassay of the HML.24  
 CC antigen or antibody is useful for diagnosis of immune disorders and  
 CC cancer, for monitoring of anti-HML.24 antibody immunotherapy, and for  
 CC assay of the antibody or antigen for investigative purposes, in  
 CC biological samples such as blood, serum, urine, milk, synovial fluid or  
 CC microorganism culture media. The method is sensitive down to 500 pg/ml  
 CC antibody.

XX  
 SQ Sequence 180 AA:

Query Match 100.0%; Score 889; DB 20; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYGKRVPMEDGDKRCKLLGIGILVLLIIVILGVPILFTTKANSEACRDLRAV 60  
 DB 1 mastsydygcrvpmedgdkrcklllgigilvllilvllgvpilfttkanseacrdglrav 60

QY 61 MECRNVTLLLOOEELTEAOGFQVDEAQAATNHNHVMILMSLDEKROGOKVLELGEI 120

QY 61 mecrnvtlllqgelteagkqfgdveaqaatcnhtvmmslmsldekeqgqkvrlelgei 120

QY 61 mecrnvtlllqgelteagkqfgdveaqaatcnhtvmmslmsldekeqgqkvrlelgei 120

QY 121 TLLNHHKLODAAEYERLRRENOYLSVRLADKKYPPSSODSSAAAPOLLIVLGLSMLLO 180

QY 121 tllnhkldgaaeevrlrrenqvlsvrladckkypssqgssaaapqlilvllglslql 180

## RESULT 6

AA02576	
ID	AA02576 standard; Protein; 180 AA
XX	

RESULT	7
AAV05484	
ID	AAV05484 standard; Protein; 180 AA
XX	
AC	AAV05484;
XX	
DT	07-JUL-1999 (first entry)

## RESULT 8

AAV07250  
ID AAV07250 standard; Protein; 180 AA

AC AAY07250;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE BST-2 protein.  
 XX  
 KW Mouse; BST-2; monoclonal antibody; RS38; myeloma; cytotoxic activity.  
 XX  
 OS Mus sp.  
 XX  
 PN JP11092399-A.  
 XX  
 PD 06-APR-1999.  
 XX  
 PE 24-SEP-1997; 97JP-0274960.  
 XX  
 PR 24-SEP-1997; 97JP-0274960.  
 XX  
 PA (CHUGAI PHARM CO LTD.  
 XX  
 DR WPI: 1999-283503/24.  
 XX  
 NR N-PSDB; AAX29996.  
 XX  
 CC This sequence represents the mouse BST-2 protein which is used to raise  
 CC antibodies, especially the monoclonal antibody RS38. The antibody can be  
 CC used in compositions to treat myelomas when the antibody is associated  
 CC with a cytotoxic activity.  
 XX  
 SO Sequence 180 AA;

Query Match 100.0%; Score 889; DB 20; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6,2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLGIGILVLLIIVLGVPLIIFITKANSEACRDLRAV 60  
 DB 1 mastsydycrvpmmedgdkrcklllgilvllivllgvplllfllkanseacrdglrav 60  
 QY 61 MECRWVTHLQOELTEAOKGFQDVEAQAATCNHTVWALMASIDAERAGQKVEELEG 120  
 DB 61 mecrwvthlqoelteaokgfgdveaqaatcnhtvwalmasidaeragqkveelegel 120  
 QY 121 TTLNHLQDASAEVERLRENOVLSVRIADKKRYPPSSQSSAAPQLLIVLLGSLALQ 180  
 DB 121 ttlnhlqdasaeverlrrenovlsvriadkkryppssqssaaapqllivllglsallq 180

RESULT 9  
 ID AAY53273  
 AC AAY53273 standard; Protein; 180 AA.  
 XX  
 AC AAY53273;  
 XX  
 DT 21-JUL-2000 (first entry)  
 XX  
 DE Human HML.24 antigen protein sequence SEQ ID NO:2.  
 XX  
 KW Human; HML.24 antigen protein; detection; plasmocytoma;  
 KW multiple myeloma; plasmocytic leukemia; extramedullary plasmocytoma;  
 KW multiple plasmocytoma; asymptomatic myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200017395-A1.  
 XX  
 PD 30-MAR-2000.  
 XX

PF 20-AUG-1999; 99WO-JP04502.  
 XX  
 PR 18-SEP-1998; 98JP-0264593.  
 XX  
 PA (CHUGAI SEIYAKU KK.  
 XX  
 PI Kawai S, Koshidara Y, Kosaka M;  
 XX  
 DR WPI: 2000-283616/24.  
 DR N-PSDB; AAA13654.  
 XX  
 PT Detection or measurement of plasmocytomas, applicable for early  
 PT diagnosis of e.g. multiple myeloma and plasmocytic leukemia, using a  
 PT polynucleotide which is expressed specifically or strongly in  
 PT plasmocytomas  
 XX  
 PS Disclosure; Page 15-16; 20pp; Japanese.  
 XX  
 CC A method has been developed for detecting or measuring plasmocytomas in  
 CC a sample at an early stage of disease development. The method comprises  
 CC amplifying a polynucleotide which is expressed specifically or strongly  
 CC in plasmocytomas before quantifying the amplification product by  
 CC comparing with results obtained with a control sample. The method is for  
 CC detecting or measuring plasmocytomas, applicable for early diagnosis of  
 CC e.g. multiple myeloma, plasmocytic leukemia, isolated plasmocytoma,  
 CC extramedullary plasmocytoma, multiple plasmocytoma obtained from smoking  
 CC or asymptomatic myeloma. The present sequence represents human HML.24  
 CC antigen protein, which is expressed in plasmocytomas and so can be used  
 CC in the method of the invention.  
 XX  
 SO Sequence 180 AA;

Query Match 100.0%; Score 889; DB 21; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6,2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLGIGILVLLIIVLGVPLIIFITKANSEACRDLRAV 60  
 DB 1 mastsydycrvpmmedgdkrcklllgilvllivllgvplllfllkanseacrdglrav 60  
 QY 61 MECRWVTHLQOELTEAOKGFQDVEAQAATCNHTVWALMASIDAERAGQKVEELEG 120  
 DB 61 mecrwvthlqoelteaokgfgdveaqaatcnhtvwalmasidaeragqkveelegel 120  
 QY 121 TTLNHLQDASAEVERLRENOVLSVRIADKKRYPPSSQSSAAPQLLIVLLGSLALQ 180  
 DB 121 ttlnhlqdasaeverlrrenovlsvriadkkryppssqssaaapqllivllglsallq 180

RESULT 10  
 ID ABB50295  
 AC ABB50295 standard; Protein; 180 AA.  
 XX  
 AC ABB50295;  
 XX  
 DT 08-FEB-2002 (first entry)  
 XX  
 DE Bone marrow stromal antigen (BST-2) ovarian tumour marker protein, #80.  
 XX  
 KW Ovarian tumour marker gene; human; overexpression; upregulation;  
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;  
 KW identification; serous cystadenoma; borderline serous tumour;  
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;  
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;  
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;  
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;  
 KW immune response pathway; cell proliferation regulation; protein folding;  
 KW membrane localised; secreted; therapeutic target; cytostatic;  
 KW gene therapy; vaccine.  
 XX  
 OS Homo sapiens.  
 XX

PN WO2001/5177-A2.  
 XX 11-OCT-2001.  
 PD  
 XX  
 PF 03-APR-2001; 2001WO-US10947.  
 XX  
 PR 03-APR-2000; 2000US-194336P.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;  
 XX  
 DR WPI: 2001-626450/72.  
 N-PSDB; ABA83121.  
 XX  
 PT Detecting and identifying ovarian tumor, identifying increased risk for  
 PT developing ovarian cancer, and determining effectiveness of ovarian  
 PT cancer treatment, by measuring expression level of ovarian tumor marker  
 gene -  
 XX  
 PS Claim 23; Page 124; 140pp; English.

CC The invention relates to methods for diagnosing and prognosing ovarian  
 CC tumors in an individual via the detection and measurement of the  
 CC expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,  
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,  
 CC ABA83181 and ABA83183). The methods of the invention are useful for  
 CC detecting an ovarian tumor in a patient, for identifying an individual  
 CC at increased risk for developing ovarian cancer, in prognostic tests for  
 CC assessing the relative severity of ovarian cancer, in tests for  
 CC monitoring a patient in remission from ovarian cancer and in tests for  
 CC monitoring disease status in a patient being treated for ovarian cancer.  
 CC The methods can additionally be used to identify a particular tumor as  
 CC being an ovarian tumor (i.e., an epithelial ovarian tumor selected from  
 CC serous cystadenoma, borderline serous tumor, serous cystadenocarcinoma,  
 CC mucinous cystadenoma, borderline mucinous tumor, mucinous  
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,  
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner  
 CC tumor). The ovarian tumor marker genes of the invention were identified  
 CC using SAGE (serial analysis of gene expression) and were found to be  
 CC overexpressed in a broad variety of ovarian epithelial tumor cells  
 CC relative to normal ovarian epithelial cells. The marker genes are  
 CC implicated in immune response pathways, in the regulation of cell  
 CC proliferation and in protein folding, and many of these are membrane-  
 CC localised or secreted. In addition to their use as diagnostic and  
 CC prognostic markers, the ovarian tumor marker genes or their encoded  
 CC proteins may be used as therapeutic targets for the treatment and  
 CC prevention of ovarian cancer. Sequences ABB50257-ABB50299 represent  
 CC proteins encoded by ovarian tumor marker genes of the invention.

XX Sequence 180 AA:  
 SQ

Query Match 100.0%; Score 889; DB 22; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6, 2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASTSYDYCVPMEDGDRCKLLGIGILVLIIVLGVPILFTTKANSEACRDGIRAV 60  
 DB 1 mastsydycvpmmedgdkrccklllgilvllivlgvplllftlkansacrdgrrav 60  
 QY 61 MECRNVTYHLLQOELTEAQKGFQDVEAQAATCNHTVMAALMASLDAEKAGQKVEELEGEI 120  
 DB 61 mecrnvtlllqelteaqqkfgdveaqaacnhtvmaalsldaeakgqkveelege1 120  
 QY 121 TLTNHLQDASAEVERLRRENOYLSVRIDAKKYYPPSSQDSSSAAAPQLLVLLGLSALIQ 180  
 DB 121 tltlnhlqdasaevertlrrenqylsvridackkyyppssqdsassaapqlllvllglisallq 180

RESULT 11  
 AAB70697  
 ID AAB70697 standard; Protein; 180 AA.

XX AAB70697;  
 AC 18-MAY-2001 (first entry)  
 XX  
 DT  
 XX  
 DE Human HML.24 protein antigen SEQ ID NO:2.  
 XX  
 KW Human; HML.24 antigen expression potentiator; HML.24 protein antigen;  
 KW myeloma; interferon alpha; interferon gamma; IFN-alpha; IFN-gamma;  
 KW interferon regulatory factor 2; IRF-2; cytostatic; cytotoxic antibody;  
 KW multiple myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200113940-A1.  
 XX  
 PD 01-MAR-2001.  
 XX  
 PF 22-AUG-2000; 2000WO-JP05617.  
 XX  
 PR 23-AUG-1999; 99JP-0236007.  
 PR 16-FEB-2000; 2000JP-0038689.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Kosaka M, Ozaki S, Wakahara Y;  
 XX  
 DR WPI: 2001-202921/20.  
 N-PSDB; AAF74792.  
 XX  
 PT HML.24 antigen expression potentiating agent containing interferon  
 PT alpha or gamma or IRF-2 for treatment of myeloma -  
 PS  
 PS Claim 1; Page 55-56; 72pp; Japanese.

XX The present invention describes an agent for potentiating the expression  
 CC of HML.24 antigen in myeloma cells. The agent contains as an active  
 CC component interferon (IFN) alpha or gamma, or interferon regulatory  
 CC factor 2 (IRF-2), or a compound promoting the expression of IRF-2. Also  
 CC described are: (1) drug compositions for the treatment of myeloma which  
 CC contain the HML.24 antigen expressions potentiating agent together with a  
 CC cytotoxic antibody binding to HML.24 antigen; (2) screening compounds  
 CC for their activity in potentiating the expression of IRF-2; and (3) kits  
 CC for the treatment of myeloma using the HML.24 antigen expression  
 CC potentiating agent and a cytotoxic antibody. The agent has cytostatic  
 CC activity and can be used for the treatment of myeloma, especially of  
 CC multiple myeloma. The present sequence represents the human HML.24  
 CC protein antigen.

XX Sequence 180 AA:  
 SQ

Query Match 100.0%; Score 889; DB 22; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 6, 2e-78;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASTSYDYCVPMEDGDRCKLLGIGILVLIIVLGVPILFTTKANSEACRDGIRAV 60  
 DB 1 mastsydycvpmmedgdkrccklllgilvllivlgvplllftlkansacrdgrrav 60  
 QY 61 MECRNVTYHLLQOELTEAQKGFQDVEAQAATCNHTVMAALMASLDAEKAGQKVEELEGEI 120  
 DB 61 mecrnvtlllqelteaqqkfgdveaqaacnhtvmaalsldaeakgqkveelege1 120  
 QY 121 TLTNHLQDASAEVERLRRENOYLSVRIDAKKYYPPSSQDSSSAAAPQLLVLLGLSALIQ 180  
 DB 121 tltlnhlqdasaevertlrrenqylsvridackkyyppssqdsassaapqlllvllglisallq 180

RESULT 12  
 AAG73947  
 ID AAG73947 standard; Protein; 193 AA.



AC AAG73947;  
XX  
DT 03-SEP-2001 (first entry)  
XX  
DE Human colon cancer antigen protein SEQ ID NO:4711.  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KM colorectal carcinoma; chromosome 19.  
XX  
OS Homo sapiens.  
XX  
PN WO200122920-A2.  
XX  
PD 05-APR-2001.  
XX  
PE 28-SEP-2000; 2000WO-US26524.  
XX  
PR 29-SEP-1999; 99US-0157137.  
PR 03-NOV-1999; 99US-0163280.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
DR WPI: 2001-235357/24.  
DR N-PSDB: AAH33378.  
XX  
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
FT useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
PS Claim 11; Page 6512-6513; 9803PP; English.  
XX  
AA AAG73943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patient's own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAH77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
SQ Sequence 193 AA:  
  
Query Match 100.0%; Score 889; DB 22; Length 193;  
Best Local Similarity 100.0%; Pred. No. 6,8e-78;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AAW77292  
ID AAW77292 standard; Protein; 180 AA.  
XX  
AC AAW77292;  
XX  
DT 14-DEC-1998 (first entry)  
XX  
DE Protein bound by Anti-HM1.24 antibody.  
XX  
DE Anti-HM1.24; antibody; lymphocyte activation inhibitor;  
KW autoimmune disease; organ transplant; allergy.  
XX  
OS Homo sapiens.  
XX  
PN WO9837913-A1.  
XX  
PD 03-SEP-1998.  
XX  
PE 27-FEB-1998; 98WO-JP00831.  
XX  
PR 28-FEB-1997; 97JP-0045663.  
XX  
PA (CHUS) CHUGAI SEIYAKU KK.  
XX  
PI Koshihara Y;  
XX  
DR WPI: 1998-480937/41.  
DR N-PSDB: AAV59114.  
XX  
PT Lymphocyte activation inhibitor comprises antibodies, particularly  
FT anti-HM1.24 antibody - for preventing and treating autoimmune  
PT diseases, rejection reactions in organ transplant or allergy  
XX  
PS Disclosure: Page 38-39; 53PP; Japanese.  
XX  
CC The Anti-HM1.24 antibody can be used in the production of lymphocyte  
CC activation inhibitors. These inhibitors can be used for the prevention  
CC and treatment of autoimmune diseases, rejection reactions in organ  
CC transplant or allergy. Administration is non-oral, e.g. by intra  
CC venous and intramuscular injection, local or systemic.  
XX  
SQ Sequence 180 AA:  
  
Query Match 99.2%; Score 882; DB 19; Length 180;  
Best Local Similarity 99.4%; Pred. No. 3e-77;  
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MASTSYDYCRVPMEDGDKRCKLLIGILVLLIIVLIGVPLITFTKANSEACRDLRAV 60  
DB 1 mastsydycrvpmegdckrcklllgilvllivlligvpllftkanskeacrdglrav 60  
QY 61 MECRNVTHTLLOELTEAOKGFODVEAQAATCNHTVMAALASLDAEKAQOKKVEELEGEL 120  
DB 61 mecrnvthlllqelteaogfodvgaqaatcnhtvmaalasaideakaqgkveelege 120  
QY 121 TFLNHLQDASAEVRLRENOVLSVRIDAKRYPPSSODSSAAAPQLIVLLIGLSALIQ 180  
DB 121 tflnhklqdasavevrlrenovlsvrldakryppssodssaaapqllivlliglsali 180

RESULT 14  
AAW36951  
ID AAW36951 standard; Protein; 197 AA.  
XX  
AC AAW36951;  
XX  
DT 12-MAY-1998 (first entry)  
XX  
DE Protein encoded by clone 0238\_1.  
XX  
KW Human; secreted protein; molecular weight marker; genetic fingerprinting;  
KW antibody production; nutritional supplement; therapy; clone 0238\_1;

KW dendritic cell.  
XX  
OS Homo sapiens.  
XX  
PN WO9740069-A2.  
XX  
PD 30-OCT-1997.  
XX  
PF 14-APR-1997; 97WO-US06134.  
XX  
PR 19-APR-1996; 96US-0635311.  
XX  
PA (GEM ) GENETICS INST INC.  
PI Jacobs K, Lavallie ER, Mccoy JM, Merberg D, Racie LA;  
PI Spaulding V;  
XX  
XX WPI: 1997-535776/49.  
XX N-PSDB: AAY00426.  
XX  
XX Isolated nucleic acid clones from ATCC 98028 encode novel secreted  
XX proteins - having many potential uses, e.g. as immunomodulators,  
XX cell proliferation or differentiation inhibitors or haematopoiesis  
XX regulators  
XX  
XX Claim 36; Page 77; 114pp; English.  
XX  
XX This sequence is a protein encoded by clone 0238.1, which is a  
XX polynucleotide of the invention. The DNA encoding this sequence was  
XX isolated from a human dendritic cell cDNA library. The polynucleotide,  
XX which encodes a secreted protein, can be used, e.g. as a tissue or  
XX molecular weight marker, in genetic fingerprinting, to raise anti-protein  
XX or anti-DNA antibodies and in interaction trap assays. The protein can be  
XX used to assay biological activity, raise antibodies for use in  
XX immunoassays, as a marker, to identify inhibitors of its interactions and  
XX as a nutritional supplement. It may also have a very wide range of  
XX therapeutic and biological activities (no examples are given to support  
XX this), e.g. cytokine or modulator of cell proliferation and  
XX differentiation, immunostimulant or immunosuppressant, haematopoiesis  
XX regulator, bone, cartilage, tendon, ligament and/or nerve tissue growth  
XX stimulator, follicle inhibitor/stimulator, chemotactic/chemokinetic,  
XX haemostatic, thrombolytic or anti-inflammatory agent, antimicrobial,  
XX biohythm, metabolism or behaviour modifier, anti-depressant or analgesic  
XX or psoriasis treatative.  
XX  
XX Sequence 197 AA;  
SQ

Query Match 94.9%; Score 843.5; DB 18; Length 197;  
Best Local Similarity 97.7%; Pred. No. 1.7e-73;  
Matches 173; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
OY 1 MASTSYDVCVPMEDSKRCKLLGIGILVLLIIVILGVPLIIFTIKANSEACRDGIRAV 60  
DB 1 mastsydvcvpmegdskrccklllgilvlllilgvplllftlkanscaacrdlirav 60  
OY 61 MCRNATHTLLOELFTEAOGKGFQVEAQAATCNHTVMAALMSLDAEKAOGKKVEELGET 120  
DB 61 mcrnathtlllgelteaogkfgdveaq-aancntvmalmsldaekagqkkveeleget 119  
OY 121 TPTLNHKLQDASAEEVERLRRENOVLSVRIDAKKRYPPSSODSSAAAPQLIVLIGLSA 177  
DB 120 tptlnhklqdasaeerlrrenqylsvrladkkyppssqssaaapqlilvlligls 176

RESULT 15  
AAY32763  
ID AAY32763 standard; Protein; 143 AA.  
XX  
XX AAY32763;  
AC  
XX  
XX 25-OCT-1999 (first entry)  
DT  
XX

DE HML.24 antigenic protein/HA peptide fusion protein sequence.  
XX  
XX HML.24; antigenic protein; antibody; immunoassay; diagnosis; cancer;  
XX immune disorder; immunotherapy monitoring.  
XX  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX FT MISC-difference 7  
XX FT MISC-difference 7 /label= encoded by GAC  
XX  
XX WO9943703-A1.  
XX  
XX 02-SEP-1999.  
XX  
XX PD  
XX  
XX PF 25-FEB-1999; 99WO-JP00885.  
XX  
XX PR 25-FEB-1998; 98JP-0060613.  
XX  
XX (CHUS ) CHUGAI SEIYAKU KK.  
XX  
XX PI Kohshihara Y, Ozaki Y;  
XX  
XX WPI, 1999-518836/43.  
XX N-PSDB: AA210915.  
XX  
XX Immunoassay of anti-HML.24 antibody or soluble HML.24 antigen,  
XX useful for diagnosis of immune disorders and cancer  
XX  
XX Disclosure; Page 77-78; 138pp; Japanese.  
XX  
XX This sequence represents a fusion protein of the human soluble HML.24  
XX antigenic protein and a HA peptide. The invention relates to an  
XX immunochemical assay of anti-HML.24 antibody by use of a soluble HML.24  
XX antigenic protein, or an immunochemical assay of the soluble antigen by  
XX use of the antibody. The immunoassay of the HML.24 antigen or antibody is  
XX useful for diagnosis of immune disorders and cancer, for monitoring of  
XX anti-HML.24 antibody immunotherapy, and for assay of the antibody or  
XX antigen for investigative purposes, in biological samples such as blood,  
XX serum, urine, milk, synovial fluid or microorganism culture media. The  
XX method is sensitive down to 500 pg/ml antibody.  
XX  
XX Sequence 143 AA;  
SQ

Query Match 73.1%; Score 649.5; DB 20; Length 143;  
Best Local Similarity 95.1%; Pred. No. 6e-55;  
Matches 135; Conservative 1; Mismatches 3; Indels 3; Gaps 1;  
OY 39 VPLIIFTIKANSEACRDGIRAVMECRNVTHTLLQDELTEAOGKGFQVEAQAATCNHTVMA 98  
DB 5 vplyagf---nseacrdgtravmecnrvthlllgelteaogkfgdveaqaatcnhtvma 1  
OY 99 MASLDAEKAOGKKVEELBEETITLNHKLQDASAEEVERLRRENOVLSVRIDAKKRYPPSSQ 158  
DB 62 masldaekagqgkveeleetltlnhklqdasaeerlrrenqylsvrladkkyppsq 121  
OY 159 DSSSAAAPQLIVLIGLSALIQ 180  
DB 122 dsssaapqlilvlliglsalig 143

Search completed: June 4, 2002, 15:26:38  
Job time: 231 sec

---

**THIS PAGE BLANK (USPTO)**